MATHEMATISCHE ZEITSCHRIFT

65. BAND

MATHEMATISCHE ZEITSCHRIFT

UNTER STÄNDIGER MITWIRKUNG VON

E. KAMKE
TÜBINGEN

K. KNOPP
TÜBINGEN

R. NEVANLINNA
HELSINKI

E. SCHMIDT
BERLIN

F. K. SCHMIDT
HEIDELBERG

HERAUSGEGEBEN VON

H. WIELANDT
TÜBINGEN

WISSENSCHAFTLICHER BEIRAT

W. BLASCHKE L. FEJÉR A. E. INGHAM
H. KNESER W. MAGNUS O. PERRON W. SÜSS

65. BAND

BERLIN · GÖTTINGEN · HEIDELBERG
SPRINGER-VERLAG
1956

Unveranderter Nachdruck 1971
Springer-Verlag Berlin Heidelberg New York

Inhalt des 65. Bandes

Seite

AVAKUMOVIĆ, V. G., Über die Eigenfunktionen auf geschlossenen Riemannschen Mannigfaltigkeiten 327

BAUER, H., Über die Beziehungen einer abstrakten Theorie des Riemann-Integrals zur Theorie Radonscher Maße 448

BOJANIĆ, R., W. JURKAT und A. PEYERIMHOFF, Über einen Taubersatz für Faltungen 195

DOLD, A., Erzeugende der Thomschen Algebra $\mathfrak{N}$ 25

DOLD, A., Vollständigkeit der Wuschen Relationen zwischen den Stiefel-Whitneyschen Zahlen differenzierbarer Mannigfaltigkeiten 200

DOLPH, C. L., and R. K. RITT, The Schwinger variational principles for one-dimensional quantum scattering 309

EICHLER, M., Berichtigung zu meiner Arbeit über den Hilbertschen Klassenkörper eines imaginärquadratischen Zahlkörpers 214

ENDL, K., Orthogonalisierung auf einem k-strahligen symmetrischen Integrationsstern 1

ENDL, K., Über eine ausgezeichnete Eigenschaft der Koeffizientenmatrizen des Laguerreschen und des Hermiteschen Polynomsystems 7

ENDL, K., Über Klassen von Limitierungsverfahren, die die Klasse der Hausdorffschen Verfahren als Spezialfall enthalten 113

FOSTER, A. L., Ideals and their Structure in Classes of Operational Algebras 70

FOULKES, H. O., Theorems of Kakeya and Pólya on power-sums 345

GRAUERT, H., und R. REMMERT, Plurisubharmonische Funktionen in komplexen Räumen 175

HUCKEMANN, F., Zur Darstellung von Riemannschen Flächen durch Streckenkomplexe 215

HUCKEMANN, F., Über den Einfluß von Randstellen Riemannscher Flächen auf die Wertverteilung 240

KRULL, W., Über reelle Radikalkörper 76

LANDSBERG, M., Lineare topologische Räume, die nicht lokalkonvex sind 104

MOHANTY, R., and S. MOHAPATRA, On the absolute logarithmic summability of a Fourier series 207

MORITA, K., and H. TACHIKAWA, Character modules, submodules of a free module, and quasi-Frobenius rings 414

NAUMANN, H., Beliebige konvexe Polytope als Schnitte und Projektionen höherdimensionaler Würfel, Simplizes und Maßpolytope 91

NEUMANN, H., On varieties of groups and their associated near-rings 36

PETERSEN, G. M., Inclusion between limitation methods 494

RAMANUJAN, M. S., Theorems on the product of quasi-Hausdorff and Abel transforms 442

Seite

SCHOENEBERG, B., Über die Diskriminante der elliptischen Funktionen und ihre Quadratwurzel 16

SCHÖNEBORN, H., Über den Zusammenhang zwischen Dualitäts- und Vollständigkeitseigenschaften bei gewissen topologischen Abelschen Gruppen 429

SCHUBERT, H., Knoten mit zwei Brücken 133

SHAH, S. M., and S. K. SINGH, On the derivative of a meromorphic function with maximum defect 171

SONNER, H., Über die Zurückführung partieller Differentialgleichungen auf gewöhnliche 483

VOGEL, W., Die Annäherung guter Strategien bei einer gewissen Klasse von Spielen 283

VOLKMANN, B., Über Hausdorffsche Dimensionen von Mengen, die durch Zifferneigenschaften charakterisiert sind. V 389

ZADDACH, A., Über Anti-Fano-Ebenen 353

ENDL, K.
Math. Zeitschr. Bd. 65, S. 1—6 (1956)

Orthogonalisierung auf einem k-strahligen symmetrischen Integrationsstern

Von
KURT ENDL

Einleitung

Im Laufe einer Untersuchung über LAGUERREsche und HERMITEsche Polynome[1]) stellte sich dem Verfasser folgendes Problem: Gibt es für beliebiges natürliches k einen Orthogonalisierungsprozeß, der Polynome von der Struktur

$$(1) \qquad P_n(z) = p_n^{(n)} \cdot z^n + p_{n-k}^{(n)} \cdot z^{n-k} + p_{n-2k}^{(n)} \cdot z^{n-2k} + \cdots \qquad (p_\nu^{(n)} = 0 \text{ für } \nu < 0)$$

erzeugt? Alle bisher bekannten Orthogonalisierungsverfahren liefern nur Polynome $P_n(z)$ mit $k = 1$ (LAGUERRE, JACOBI) und $k = 2$ (HERMITE, GEGENBAUER, LEGENDRE, TSCHEBYSCHEFF). *Es wird gezeigt, daß es für jedes natürliche k Orthogonalisierungsverfahren gibt, die Polynome von der Bauart* (1) *erzeugen.* Das wesentliche Kennzeichen jedes dieser Verfahren besteht darin, in bezug auf Orthogonalisierungs-„Intervall" und Belegungsfunktion eine Symmetrie der Ordnung k zu enthalten: Es wird auf einem k-strahligen symmetrischen Stern mit Hilfe einer Belegungsfunktion orthogonalisiert, die invariant gegen solche Drehungen des Sternes ist, die ihn in sich selbst überführen. Zum Schlusse wird bewiesen, daß jedes auf einem k-strahligen Stern orthogonales Polynomsystem $\{P_n(z)\}$ k Teilsysteme $\{P_m^{(\varrho)}(z)\}$ $(\varrho = 0, 1, \ldots, k-1)$ enthält, von denen jedes für sich im klassischen Sinn ein orthogonales Polynomsystem bildet, und daß umgekehrt durch „Ineinanderschachteln" von k geeigneten (im klassischen Sinne orthogonalen) Polynomsystemen, ein neues Polynomsystem erzeugt werden kann, das orthogonal auf einem k-strahligen Stern ist, und dessen Polynome die Struktur (1) besitzen.

§ 1. Der allgemeine Orthogonalisierungsprozeß und die Lückeneigenschaft der zugehörigen orthogonalen Polynome

Die Ursache dafür, daß bei Orthogonalisierung der Potenzen x^n auf einem Intervall $[a, b]$ mit einer Belegungsfunktion $p(x)$ Polynome von der Struktur (1) mit $k = 2$ auftreten (also abwechselnd gerade und ungerade Polynome) liegt darin, daß die Belegungsfunktion $p(x)$ in bezug auf das Orthogonalisierungsintervall $[a, b]$ eine Symmetrieeigenschaft besitzt[2]). Diese einfache

[1]) ENDL, KURT: Sur une classe de polynômes orthogonaux généralisant ceux de Laguerre et Hermite. C. R. Acad. Sci. **241**, 723—724 (1955) und ENDL, KURT: Über eine ausgezeichnete Eigenschaft der Koeffizientenmatrizen des LAGUERREschen und des HERMITEschen Polynomsystems. Math. Z. **65**, 7—15 (1956).

[2]) Dies folgt durch eine einfache Überlegung aus dem ERHARD-SCHMIDTschen Orthogonalisierungsverfahren.

Bemerkung läßt vermuten, daß ein Orthogonalisierungsprozeß, der für beliebiges natürliches k Polynome von der Form (1) erzeugen soll, eine Symmetrie in bezug auf Belegungsfunktion und Orthogonalisierungs-„Intervall" besitzen muß, wobei diese Symmetrie noch in irgendeiner Form den Parameter k enthalten muß. Wir betrachten dazu in der komplexen z-Ebene k Strecken $\mathfrak{S}_\lambda$ in symmetrischer Anordnung (k fest, $\lambda = 0, 1, \ldots, k-1$), die im 0-Punkt zusammenlaufen; es sei z.B. $\mathfrak{S}_\lambda = \mathfrak{S}_\lambda^{(R,k)}$ die Menge der Punkte $z = r \cdot e^{\frac{2\pi i}{k}\lambda}$ $(0 \leq r \leq R)$. Diese k Strecken $\mathfrak{S}_\lambda$ bilden eine Punktmenge, die wir einen k-strahligen symmetrischen Stern nennen und mit $\mathfrak{S} = \mathfrak{S}_R^{(k)}$ bezeichnen. Ferner wählen wir eine Belegungsfunktion $p(z)$, die invariant gegen Drehungen um $\frac{2\pi}{k}$ ist, für die also gilt $p(z) = p\left(z \cdot e^{\frac{2\pi i}{k}}\right)$. Definieren wir nun als skalares Produkt

$$(2) \qquad (P, Q)^{(k)}_{p(z)} = (P, Q) = \int_{\mathfrak{S}} P \overline{Q}\, p\, dr = \sum_{\lambda=0}^{k-1} \int_{\mathfrak{S}_\lambda}^{R} P(z)\, \overline{Q(z)}\, p(z)\, dr,$$

so liegt offenbar ein Orthogonalisierungsprozeß vor, der in bezug auf Belegungsfunktion und „Orthogonalisierungsintervall" eine Symmetrieeigenschaft besitzt. Durch Orthogonalisierung der Potenzen $\{z^n\}$ mit Hilfe des skalaren Produktes (2) entsteht nun ein orthogonales Polynomsystem

$$P_n(z) = p_n^{(n)} \cdot z^n + p_{n-1}^{(n)} z^{n-1} + \cdots + p_0,$$

das wir uns gleich so normiert denken, daß $p_n^{(n)} = 1$ ist $(n = 0, 1, \ldots)$. Wir behaupten folgenden

Satz I. *Alle Polynome $P_n(z)$ besitzen die Struktur* (1)

$$P_n(z) = 1 \cdot z^n + p_{n-k}^{(n)} \cdot z^{n-k} + p_{n-2k}^{(n)} \cdot z^{n-2k} + \cdots,$$

d.h. $p_\nu^{(n)} = 0$ für $\nu \not\equiv n \pmod{k}$ $(0 \leq \nu \leq n = 0, 1, \ldots)$. Nimmt weiterhin die Belegungsfunktion $p(z)$ auf einem Strahl $\mathfrak{S}_\lambda$ nur reelle Werte an, so sind die Polynome $P_n(z)$ reell $(n = 0, 1, \ldots)$.

Beweis. $P_0(z) \equiv p_0^{(0)} = 1$ hat trivialerweise die Form (1). Wir nehmen an, dies sei auch für alle $P_\nu(z)$ mit $0 \leq \nu < n$ erfüllt, d.h.

$$P_\nu(z) = z^\nu + p_{\nu-k}^{(\nu)} z^{\nu-k} + \cdots \qquad (p_\mu^{(\nu)} = 0 \text{ für } \mu < 0)$$

oder wenn wir alle natürlichen Zahlen n in der Form zerlegen $n = n_k^* + n_k \cdot k$ $(0 \leq n_k^* < k)$

$$(3) \qquad P_\nu(z) = \sum_{\varkappa=0}^{\nu_k^*} p_{\nu_k^* + \varkappa k}^{(\nu)} \cdot z^{\nu_k^* + \varkappa k}.$$

Nach ERHARD SCHMIDT erhält man für $P_n(z)$ die Beziehung

$$(4) \qquad P_n(z) = z^n - \sum_{\nu=0}^{n-1} C_\nu P_\nu(z),$$

wobei $C_\nu = \frac{(z^n, P_\nu)}{(P_\nu, P_\nu)}$ ist.

Für $(z^n, P_\nu(z))$ folgt aus (3)

$$(z^n, P_\nu) = \sum_{\lambda=0}^{k-1} \int_{\mathfrak{S}_\lambda}^{R} z^n \overline{P_\nu(z)}\, p(r)\, dr$$

$$= \sum_{\varkappa=0}^{\nu_k} \overline{p_{\nu_k^*+\varkappa k}^{(\nu)}} \cdot \left\{ \sum_{\lambda=0}^{k-1} \int_{\mathfrak{S}_\lambda}^{R} z^n \cdot \overline{z^{\nu_k^*+\varkappa k}} \cdot p(r)\, dr \right\}$$

$$= \sum_{\varkappa=0}^{\nu_k} \overline{p_{\nu_k^*+\varkappa k}^{(\nu)}} \cdot \int_0^R r^{n+\nu_k^*+\varkappa k} \left\{ \sum_{\lambda=0}^{k-1} \left[e^{\frac{2\pi i}{k}(n-\nu_k^*-\varkappa k)} \right]^\lambda \right\} p(r)\, dr.$$

Ist $\nu \equiv n(k)$, d.h. $\nu_k^* = n_k^*$, so ist $n - \nu_k^* - \varkappa k = (n_k - \varkappa) \cdot k \equiv o(k)$ d.h.

$$\sum_{\lambda=0}^{k-1} \left[e^{\frac{2\pi i}{k}(n-\nu_k^*-\varkappa k)} \right]^\lambda = \sum_{\lambda=0}^{k-1} 1 = k$$

und damit

$$(z^n, P_\nu) = k \cdot \sum_{\varkappa=0}^{\nu_k} \overline{p_{\nu_k^*+\varkappa k}^{(\nu)}} \cdot \int_0^R r^{n+\nu_k^*+\varkappa k}\, p(r)\, dr.$$

Dagegen gilt für $\nu \not\equiv n(k)$ wegen $e^{\frac{2\pi i}{k}(n-\nu_k^*-\varkappa k)} \neq 1$

$$\sum_{\lambda=0}^{k-1} \left[e^{\frac{2\pi i}{k}(n-\nu_k^*-\varkappa k)} \right]^\lambda = \frac{e^{\frac{2\pi i}{k}(n-\nu_k^*-\varkappa k)\,k} - 1}{e^{\frac{2\pi i}{k}(n-\nu_k^*-\varkappa k)} - 1} = 0$$

und damit $(z^n, P_\nu) = 0$.

In der Entwicklung (4) treten also nur Koeffizienten C_ν mit $\nu \equiv n(k)$ auf. Wir können schreiben:

$$P_n(z) = z^n - \sum_{\nu=0}^{n_k-1} C_{n_k^*+\nu k} \cdot P_{n_k^*+\nu k}(z)$$

$$= z^n - \sum_{\nu=0}^{n_k-1} C_{n_k^*+\nu k} \cdot \sum_{\varkappa=0}^{\nu} p_{n_k^*+\varkappa k}^{(n_k^*+\nu k)} \cdot z^{n_k^*+\varkappa k},$$

woraus nach Vertauschung der Summationen folgt:

$$(5) \qquad P_n(z) = z^n - \sum_{\varkappa=0}^{n_k-1} z^{n_k^*+\varkappa k} \cdot \left\{ \sum_{\nu=\varkappa}^{n_k-1} C_{n_k^*+\nu k} \cdot p_{n_k^*+\varkappa k}^{(n_k^*+\nu k)} \right\}.$$

Hiermit ist der erste Teil des Satzes I bewiesen. Um den zweiten Teil nachzuprüfen, sei als Induktionsannahme noch hinzugenommen, daß alle Polynome $P_\nu(z)$ reelle Polynome sind. Dann sind, da $\int_0^R r^{n+\nu_k^*+\varkappa k} p(r)\, dr$ reell ist, C_ν und damit alle Koeffizienten in (5) reell.

§ 2. Die orthogonalen Untersysteme des Systemes $\{P_n(z)\}$

Wie an anderer Stelle gezeigt wurde[3]), läßt sich mit Hilfe des in § 1 wiedergegebenen Orthogonalisierungsverfahrens [man setzt in (2) $R=\infty$ und $p(z)=e^{-z^k}$] ein mit einem natürlichen Parameter k versehenes Polynomsystem $\{P_{k,n}(z)\}$ erzeugen, das als Spezialfälle die LAGUERREschen Polynome $\{L_n(z)\}$ und HERMITEschen Polynome $\{H_n(z)\}$ umfaßt. Es gilt $\{P_{1,n}\}=\{L_n\}$ und $\{P_{2,n}\}=\{H_n\}$. Die Polynome $P_{k,n}(z)$ lauten explizit

$$P_{k,n}(z)=n_k!\sum_{\nu=0}^{n_k}\frac{(-)^{n_k-\nu}}{\nu!}\binom{n_k+\dfrac{2n_k^*+1}{k}-1}{n_k-\nu}z^{n_k^*+\nu k}.$$

Vergleicht man diesen Ausdruck mit dem der verallgemeinerten LAGUERREschen Polynome

$$L_n^{(\alpha)}(z)=n!\sum_{\nu=0}^{n}(-)^{n-\nu}\binom{n+\alpha}{n-\nu}\frac{z^\nu}{\nu!},$$

so bemerkt man eine gewisse Verwandtschaft. In der Tat lassen sich die Polynome $P_{k,n}(z)$ durch die Polynome $L_n^{(\alpha)}(z)$ folgendermaßen ausdrücken:

$$P_{k,n}(z)=z^{n_k^*}\cdot L_{n_k}^{\left(\frac{2n_k^*+1}{k}-1\right)}(z^k).$$

Der Rest n_k^* geht hierbei in den Parameterwert α ein, das Vielfache n_k in den Grad des LAGUERREschen Polynomes. Betrachten wir für den Augenblick die Unterklasse der Polynome $P_{k,n}(z)$ mit demselben Rest ϱ $(0, 1, \ldots, k-1)$, d.h. die Polynome $\{P_{k,\varrho+mk}(z)\}$ $(m=0, 1, \ldots)$, so finden wir: Das Untersystem $\{P_{k,\varrho+mk}(z)\}$ ist bis auf Multiplikation mit der Potenz z^ϱ und bis auf die Substitution $z^k=\zeta$ ein im klassischen Sinn orthogonales Polynomsystem. Es ist nämlich orthogonal auf dem Intervall $[0, \infty]$ mit der Belegungsfunktion $\zeta^{\frac{2\varrho+1}{k}-1}e^{-\zeta}$. Da dies für $\varrho=0, 1, \ldots, k-1$ gilt, folgt die Aussage: Das System $\{P_{k,n}(z)\}$ ist aus k im klassischen Sinn orthogonalen Polynomsystemen zusammensetzt.

Dieser Sachverhalt, den wir in diesem Spezialfall bemerkt haben, ist allgemein richtig[4]). Jedes auf einem k-strahligen Stern orthogonale Polynomsystem entsteht durch Ineinanderschachtelung von k (im klassischen Sinn orthogonalen) Polynomsystemen; umgekehrt kann man k geeignete (klassische) Systeme zu einem neuen System zusammensetzen. Um dies zu präzisieren,

---

[3]) ENDL, KURT: Les polynomes de Laguerre et de Hermite comme cas particuliers d'une classe de polynomes orthogonaux. Ann. Ec. Norm. (3), 73, fasc. 1, 1—13 (1956).

[4]) Der Verfasser glaubte ursprünglich, daß es sich um eine besondere Eigentümlichkeit des Systems $P_{k,n}(z)$ handele, bedingt durch die Belegungsfunktion e^{-z^k}. Daß dieser Sachverhalt allgemein zutreffend ist, geht auf eine Diskussionsbemerkung von H. WOLFGANG HAHN im Anschluß an den Vortrag des Verfassers auf der DMV-Tagung 1955 in Göttingen zurück.

betrachten wir das auf dem k-strahligen Stern $\mathfrak{S}$ mit der Belegungsfunktion $p(z)$ orthogonale Polynomsystem

$$P_n(z) = \sum_{\nu=0}^{n} p_\nu^{(n)} z^\nu.$$

Nach § 1 können wir $P_n(z)$ in der Form schreiben:

$$P_n(z) = P_{n_k^* + n_k \cdot k}(z) = \sum_{\nu=0}^{n_k} p_{n_k^*+\nu k}^{(n)} z^{n_k^*+\nu k} = z^{n_k^*} \cdot \sum_{\nu=0}^{n_k} p_{n_k^*+\nu k}^{(n)} (z^k)^\nu.$$

Wir betrachten nun alle Polynome $P_n(z)$ in derselben Restklasse ϱ ($\varrho = 0, 1, \ldots, k-1$)

$$P_{\varrho+mk}(z) = z^\varrho \sum_{\nu=0}^{m} p_{\varrho+\nu k}^{(\varrho+mk)} (z^k)^\nu \qquad (\varrho \text{ fest}, \ m = 0, 1, \ldots).$$

Diese Polynome bilden ein Untersystem des Gesamtsystems $\{P_n(z)\}$. Wenn wir von der multiplikativen Potenz z^ϱ absehen und $z^k = \zeta$ setzen, ist dieses Untersystem ein vollständiges Polynomsystem im gewöhnlichen Sinn, d.h. es enthält Polynome aller Grade m. Wir bezeichnen die Polynome des so erhaltenen ϱ-ten Polynomsystems mit

$$P_m^{(\varrho)}(\zeta) = \sum_{\nu=0}^{m} p_{\varrho+\nu k}^{(\varrho+mk)} \zeta^\nu$$

und behaupten

Satz II. A. *Jedes der Systeme $\{P_m^{(\varrho)}(\zeta)\}$ ist schon für sich ein im klassischen Sinn orthogonales Polynomsystem. Es ist orthogonal auf dem Intervall $[0, R^k]$ mit der Belegungsfunktion $p_\varrho(\zeta) = \zeta^{\frac{2\varrho+1}{k}-1} \cdot p\left(\zeta^{\frac{1}{k}}\right)$:*

$$(P_n^{(\varrho)}, P_m^{(\varrho)})_\varrho = \int_0^{R^k} P_n^{(\varrho)}(\zeta)\, \overline{P_m^{(\varrho)}(\zeta)}\, p_\varrho(\zeta)\, d\zeta = 0 \qquad \text{für } n \neq m.$$

Hierzu gilt die folgende Umkehrung:

B. *Ist k eine beliebige natürliche Zahl, $p(\zeta)$ eine beliebige Belegungsfunktion auf dem Intervall $[0, R^k]$, und bezeichnen wir mit $\{P_m^{(\varrho)}(\zeta)\}$ das zu $[0, R^k]$ und der Belegungsfunktion $p_\varrho(\zeta) = \zeta^{\frac{2\varrho+1}{k}-1} p(\zeta)$ gehörige Polynomsystem*[5] ($\varrho = 0, 1, \ldots, k-1$), *so lassen sich diese k Polynomsysteme durch die Festsetzung*

$$P_n(z) = z^{n_k^*} P_{n_k^*}^{(n_k^*)}(z^k)$$

zu einem neuen Polynomsystem $\{P_n(z)\}$ zusammensetzen. Dieses neue System $\{P_n(z)\}$ ist orthogonal auf dem k-strahligen Stern $\mathfrak{S}$ mit der Belegungsfunktion $p(z^k)$.

[5]) Falls die Belegungsfunktion $p(\zeta)$ nicht reell ist, muß hierbei das skalare Produkt lauten:

$$(P, Q)_\varrho = \int P(\zeta)\, \overline{Q(\zeta)}\, p_\varrho(\zeta)\, d\zeta.$$

Wir beginnen mit dem Beweis des Teiles A. Wegen der Orthogonalität der Polynome $P_n(z)$ auf $\mathfrak{S}$ gilt:

$$(6)\quad \left(z^\varrho P_n^{(\varrho)}(z^k), z^\varrho P_m^{(\varrho)}(z^k)\right)^{(k)}_{p(z)} = \sum_{\lambda=0}^{k-1} \int_{\mathfrak{S}_\lambda} \int_0^R z^\varrho P_n^{(\varrho)}(z^k) \cdot \overline{z^\varrho P_n^{(\varrho)}(z^k)}\, p(z)\, dr = 0 \quad \text{für } m \neq n.$$

Dieses skalare Produkt formen wir um

$$\sum_{\lambda=0}^{k-1} \int_{\mathfrak{S}_\lambda} \int_0^R z^\varrho P_n^{(\varrho)}(z^k)\, \overline{z^\varrho P_m^{(\varrho)}(z^k)}\, p(r)\, dr -$$

$$= \sum_{\nu=0}^{n} \sum_{\mu=0}^{m} p_{\varrho+\nu k}^{(\varrho+n k)} \cdot \overline{p_{\varrho+\mu k}^{(\varrho+m k)}} \int_0^R r^{2\varrho+\nu k+\mu k} \left\{ \sum_{\lambda=0}^{k-1} \left[e^{\frac{2\pi i}{k}(\varrho+\nu k-\varrho-\mu k)} \right]^\lambda \right\} p(r)\, dr$$

$$= k \int_0^R \left(\sum_{\nu=0}^{n} p_{\varrho+\nu k}^{(\varrho+n k)} (r^k)^\nu \right) \left(\sum_{\mu=0}^{m} p_{\varrho+\mu k}^{(\varrho+m k)} (r^k)^\mu \right) r^{2\varrho}\, p(r)\, dr$$

$$= k \int_0^R P_n^{(\varrho)}(r^k) \cdot P_m^{(\varrho)}(r^k)\, r^{2\varrho}\, p(r)\, dr.$$

Die Substitution $r^k = \zeta$ liefert weiter

$$= \int_0^{R^k} P_n^{(\varrho)}(\zeta)\, P_m^{(\varrho)}(\zeta)\, \zeta^{\frac{2\varrho+1}{k}-1}\, p\left(\zeta^{\frac{1}{k}}\right) d\zeta,$$

woraus nach (6) für $m \neq n$ folgt

$$(P_n^{(\varrho)}, P_m^{(\varrho)})_\varrho = 0 \qquad \text{w. z. b. w.}$$

Zum Beweis von B hat man nur die eben durchgeführten Überlegungen rückwärts zu verfolgen.

Gießen, Mathematisches Institut der Universität

(Eingegangen am 9. September 1955)

ENDL, K.
Math. Zeitschr. Bd. 65, S. 7—15 (1956)

Über eine ausgezeichnete Eigenschaft der Koeffizientenmatrizen des Laguerreschen und des Hermiteschen Polynomsystems

Von

KURT ENDL

§ 1. Einleitung

Bei Untersuchungen über Zusammenhänge zwischen LAGUERREschen und HERMITEschen Polynomen[1]) bemerkte der Verfasser eine eigenartige Eigenschaft der LAGUERREschen Polynome[2])

$$L_n(x) = n! \sum_{\nu=0}^{n} \frac{(-)^{n-\nu}}{\nu!} \binom{n}{\nu} x^\nu .$$

Diese Polynome lassen sich so normieren, d.h. jedes Polynom $L_n(x)$ läßt sich so mit einem Faktor c_n versehen, daß die zu dem neuen System $L_n^* = c_n L_n$ gehörige Koeffizientenmatrix $\mathfrak{L}^*$ der Matrizengleichung

(1) $\qquad \mathfrak{L}^* = (\mathfrak{L}^*)^{-1}$ oder $(\mathfrak{L}^*)^2 = \mathfrak{E}$ ($\mathfrak{E}$ Einheitsmatrix)

genügt. Dies kann sogar auf zwei Arten geschehen. In einer dieser ausgezeichneten Normierungen lauten die LAGUERREschen Polynome:

$$L_n^*(x) = n! \sum_{\nu=0}^{n} \frac{(-)^\nu}{\nu!} \binom{n}{\nu} x^\nu .$$

Die Beziehung (1) ist gleichbedeutend mit der Tatsache, daß die Entwicklung der Potenz x^n nach den LAGUERREschen Polynomen $L_\nu^*(x)$ $(0 \leq \nu \leq n)$ dieselben Koeffizienten besitzt wie das Polynom $L_n^*(x)$ selbst:

$$x^n = n! \sum_{\nu=0}^{n} \frac{(-)^\nu}{\nu!} \binom{n}{\nu} L_\nu^*(x) .$$

---

[1]) ENDL, KURT: Sur les systèmes de polynômes orthogonaux en involution. C. R. Acad. Sci. **241**, 682—684 (1955).

[2]) Für die gesamte Arbeit vgl. z. B. MAGNUS-OBERHETTINGER, Formeln und Sätze für die speziellen Funktionen der mathematischen Physik. Wir normieren jedes orthogonale Polynomsystem $\{P_n(x)\}$ in der Form:

$$\{P_n(x)\} = \left\{ \sum_{\nu=0}^{n} p_\nu^{(n)} x^\nu \right\} = \left\{ x^n + \sum_{\nu=0}^{n-1} p_\nu^{(n)} x^\nu \right\} (n = 0, 1, \ldots).$$

Diese einfache Bemerkung wirft das folgende Problem auf: Handelt es sich hier um eine bisher unbemerkte allgemeine Eigenschaft orthogonaler Polynomsysteme, d.h. kann man jedes orthogonale Polynomsystem

$$\{P_n(x)\} = \left\{ \sum_{\nu=0}^{n} p_\nu^{(n)} x^\nu \right\} \qquad (p_n^{(n)} = 1)$$

mit einer solchen Normierung $\{c_n\}$ versehen, daß die zu dem neuen System $\{P_n^*(x)\} = \{c_n P_n(x)\}$ gehörige Koeffizientenmatrix $\mathfrak{P}^* = (c_n p_\nu^{(n)})$ der zu (1) analogen Beziehung genügt:

$$\mathfrak{P}^* = (\mathfrak{P}^*)^{-1} \quad \text{oder} \quad (\mathfrak{P}^*)^2 = \mathfrak{E}. \tag{2}$$

Dies ist nun nicht der Fall. Nennen wir ein orthogonales Polynomsystem $\{P_n(x)\}$, für das eine solche Normierung gefunden werden kann, kurz ein involutorisches Polynomsystem, und das System $\{P_n^*(x)\} = \{c_n P_n(x)\}$, für dessen Matrix $\mathfrak{P}^*$ (2) erfüllt ist, eine involutorische Form des Systems $\{P_n(x)\}$, so gilt:

Das LAGUERREsche *System* $\{L_n^{(\alpha)}(x)\}$ *und das* HERMITEsche *System* $\{H_n(x)\}$ *sind unter den klassischen orthogonalen Polynomsystemen die einzigen involutorischen Systeme.*

Dieses Ergebnis, das auf eine außerordentliche strukturelle Verwandtschaft der LAGUERREschen und HERMITEschen Polynome hinweist, war der Anlaß zu der folgenden Fragestellung: Kann man aus dieser Verwandtschaft der beiden Systeme $\{L_n\}$ und $\{H_n\}$ nicht schon auf die Existenz eines allgemeineren übergeordneten orthogonalen Polynomsystems schließen, das die Systeme $\{L_n\}$ und $\{H_n\}$ als Spezialfälle enthält. Der Verfasser hat mit Hilfe eines von ihm eingeführten neuen Orthogonalisierungsverfahrens[3]) ein von einem ganzzahligen Parameter k ($k = 1, 2, 3, \ldots$) abhängiges orthogonales Polynomsystem $\{P_{k,n}(x)\}$ konstruiert, das für spezielle Parameterwerte das LAGUERREsche und HERMITEsche Polynomsystem enthält[4]). Es gilt:

$$\{P_{1,n}(x)\} = \{L_n(x)\}, \qquad \{P_{2,n}(x)\} = \{H_n(x)\}.$$

Dies sei in diesem Zusammenhang deswegen erwähnt, weil auch für alle $k > 2$ das System $\{P_{k,n}(x)\}$ involutorisch ist.

§ 2. Notwendige und hinreichende Bedingungen für die involutorische Eigenschaft eines orthogonalen Polynomsystems

Wir betrachten ein beliebiges orthogonales Polynomsystem.

$$\{P_n(x)\} = \left\{ \sum_{\nu=0}^{n} p_\nu^{(n)} x^\nu \right\} \qquad (p_n^{(n)} = 1;\ n = 0, 1, \ldots)$$

[3]) ENDL, KURT: Orthogonalisierung auf einem k-strahligen symmetrischen Integrationsstern. Math. Z. **65**, 1—6 (1956).

[4]) ENDL, KURT: Les Polynomes de Laguerre et de Hermite comme cas particuliers d'une classe de polynomes orthogonaux. Ann. Ec. Norm. (3), **73**, fasc. 1, 1—13 (1956) und ENDL, KURT: Sur une classe de polynomes orthogonaux généralisant ceux de Laguerre et Hermite. C. R. Acad. Sci. **241**, 723—724 (1955).

und bezeichnen wie üblich mit (P_m, P_n) das skalare Produkt, mit dessen Hilfe man durch Orthogonalisation der Potenzen z^n das System $\{P_n(x)\}$ erhält. Zu jedem solchen System gehört eine unendliche Dreiecksmatrix, die Koeffizientenmatrix $\mathfrak{P} = (p_\nu^{(n)})$. Ebenso betrachten wir die inversen Formeln, d.h. die Entwicklungen der Potenzen x^n nach den Polynomen $\{P_n\}$

$$x^n = \sum_{\nu=0}^{n} q_\nu^{(n)} P_\nu(x).$$

Diese Entwicklungen liefern uns eine neue Matrix $(q_\nu^{(n)})$, die wir mit $\mathfrak{P}^{-1}$ bezeichnen; denn es gilt:

$$(q_\nu^{(n)})\,(p_\nu^{(n)}) = (p_\nu^{(n)})\,(q_\nu^{(n)}) = \mathfrak{E} \qquad (\mathfrak{E} \text{ Einheitsmatrix}).$$

Die Frage, welche Polynomsysteme $\{P_n(x)\}$ involutorisch sind, ist gleichbedeutend mit der folgenden: Für welche Systeme $\{P_n(x)\}$ kann man Koeffizienten $c_n \neq 0$ $(n = 0, 1, \ldots)$ so finden, daß für die Polynome

$$P_n^*(x) = c_n \cdot P_n(x) = \sum_{\nu=0}^{n} c_n\, p_\nu^{(n)} x^\nu = \sum_{\nu=0}^{n} p_\nu^{*(n)} x^\nu$$

gilt:

$$x^n = \sum_{\nu=0}^{n} p_\nu^{*(n)} P_\nu^*(x) \qquad (n = 0, 1, 2, \ldots). \tag{3}$$

Nehmen wir an, für $\{P_n(x)\}$ seien Koeffizienten c_n gefunden die (3) erfüllen. Da die Koeffizienten $a_\nu^{(n)}$ in der Entwicklung

$$x^n = \sum_{\nu=0}^{n} a_\nu^{(n)} P_\nu^*(x)$$

bekanntlich lauten:

$$a_\nu^{(n)} = \frac{(x^n, P_\nu^*)}{(P_\nu^*, P_\nu^*)} = \frac{1}{c_\nu} \cdot \frac{(x^n, P_\nu)}{(P_\nu, P_\nu)} \qquad 0 \leq \nu \leq n,$$

folgt aus (3) die Aussage:

$$a_\nu^{(n)} = p_\nu^{*(n)} = c_n \cdot p_\nu^{(n)} = \frac{1}{c_\nu} \cdot \frac{(x^n, P_\nu)}{(P_\nu, P_\nu)} \qquad 0 \leq \nu \leq n = 0, 1, \ldots$$

oder

$$c_n \cdot c_\nu \cdot p_\nu^{(n)} = \frac{(x^n, P_\nu)}{(P_\nu, P_\nu)} \qquad 0 \leq \nu \leq n = 0, 1, \ldots. \tag{4}$$

Ist umgekehrt (4) erfüllt, so folgt (3). Setzen wir in (4) noch $\nu = n$, so erhalten wir wegen $p_n^{(n)} = 1$ und der Beziehung

$$\left(\sum_{\nu=0}^{n-1} p_\nu^{(n)} x^\nu, P_n\right) = 0$$

$$c_n^2 = \frac{(x^n, P_n)}{(P_n, P_n)} = \frac{\left(x^n + \sum_{\nu=0}^{n-1} p_\nu^{(n)} x^\nu, P_n\right)}{(P_n, P_n)} = \frac{(P_n, P_n)}{(P_n, P_n)} = 1,$$

d.h. $c_n = \pm 1$ $(n = 0, 1, \ldots)$. Damit gilt der

Satz I. *Das orthogonale Polynomsystem* $\{P_n(x)\} = \left\{\sum_{\nu=0}^{n} p_\nu^{(n)} x^\nu\right\}$ $(p_n^{(n)} = 1)$ *ist genau dann involutorisch, wenn sich ein Vorzeichensystem* $\{c_n\}$ *so finden läßt, daß für alle Wertepaare* (n, ν) *mit* $0 \leq \nu \leq n = 0, 1, \ldots$ *die Beziehung erfüllt ist:*

$$c_n \cdot c_\nu \cdot p_\nu^{(n)} = \frac{(x^n, P_\nu)}{(P_\nu, P_\nu)}. \tag{5}$$

Nützlich für die Anwendung in Einzelfällen ist das folgende *Korollar:* Das orthogonale Polynomsystem $\{P_n(x)\}$ ist bestimmt nicht involutorisch, wenn auch nur für ein Wertepaar (n, ν) $|p_\nu^{(n)}| \neq \left|\frac{(x^n, P_\nu)}{(P_\nu, P_\nu)}\right|$ ist.

§ 3. Behandlung der klassischen Polynomsysteme

Beginnen wir mit den LAGUERREschen Polynomen

$$L_n^{(\alpha)}(x) = n! \sum_{\nu=0}^{n} \frac{(-)^{n-\nu}}{\nu!} \binom{n+\alpha}{n-\nu} x^\nu.$$

Man berechnet leicht:

$$(x^n, L_\nu^{(\alpha)}) = \int_0^\infty x^n L_\nu^{(\alpha)} x^\alpha e^{-x}\, dx$$

$$= \Gamma(n+\alpha+1) \cdot \nu! \binom{n}{\nu}.$$

Da $(L_\nu^{(\alpha)}, L_\nu^{(\alpha)}) = \Gamma(1+\alpha)\,(\nu!)^2 \binom{\nu+\alpha}{\nu}$ ist, ergibt (5):

$$c_n \cdot c_\nu = (-)^n (-)^\nu.$$

Diese Gleichung erlaubt zwei Lösungssysteme $\{c_n^{(1)}\} = \{(-)^n\}$ und $\{c_n^{(1)}\} = \{(-)^{n+1}\}$. Daher folgt:

Das LAGUERRE*sche System ist involutorisch. Es besitzt genau zwei involutorische Formen.*

Das System $L_n^*(x)$ von § 1 erhält man offenbar für $\{c_n^{(1)}\} = \{(-)^n\}$ und $\alpha = 0$. Allgemein gilt mit den Vorzeichen $\{c_n^{(1)}\}$

$$L_n^{*(\alpha)}(x) = n! \sum_{\nu=0}^{n} \frac{(-)^\nu}{\nu!} \binom{n+\alpha}{n-\nu} x^\nu$$

und daher:

$$x^n = n! \sum_{\nu=0}^{n} \frac{(-)^\nu}{\nu!} \binom{n+\alpha}{n-\nu} L_\nu^{*(\alpha)}(x).$$

Zur Behandlung der HERMITEschen Polynome

$$H_n(x) = \frac{1}{2^n} \sum_{\nu=0}^{[n/2]} (-)^\nu \binom{n}{2\nu} \frac{(2\nu)!}{\nu!} (2x)^{n-2\nu}$$

ersetzen wir in (5) ν durch $n-2\nu$. Da durch eine ähnliche Rechnung wie für die LAGUERREschen Polynome folgt:

$$(x^n, H_{n-2\nu}) = \int_{-\infty}^{+\infty} x^n H_{n-2\nu} e^{-x^2} dx = \frac{1}{2^n} \cdot \frac{n!}{\nu!} \sqrt{\pi}$$

und da $(H_{n-2\nu}, H_{n-2\nu}) = \frac{(n-2\nu)!}{2^{n-2\nu}} \cdot \sqrt{\pi}$ ist, folgt aus (5) als Bestimmungsgleichung für die Vorzeichen c_n:

$$c_n \cdot c_{n-2\nu} = (-)^\nu \qquad 0 \leq \nu \leq \left[\frac{n}{2}\right];\ n = 0, 1, 2, \ldots .$$

Setzen wir hier $\nu = [n/2]$, so ergibt sich für $n = 2m$: $c_{2m} = (-)^m c_0$ für $n = 2m+1$: $c_{2m+1} = (-)^m c_1$. Es können also c_0 und c_1 frei gewählt werden. Wir erhalten vier Vorzeichensysteme:

	c_0	c_1	c_2	c_3	c_4	c_5	c_6	c_7	...
$\{c_n^{(1)}\}$	+	+	−	−	+	+	−	−	...
$\{c_n^{(2)}\}$	+	−	−	+	+	−	−	+	...
$\{c_n^{(3)}\}$	−	+	+	−	−	+	+	−	...
$\{c_n^{(4)}\}$	−	−	+	+	−	−	+	+	... .

Es gilt also: *Das* HERMITE*sche System ist involutorisch. Es besitzt genau vier involutorische Formen.*

Die Vorzeichen $\{c_n^{(1)}\}$ werden durch $\{(-)^{[n/2]}\}$ gegeben. Die zugehörige involutorische Form der HERMITEschen Polynome lautet:

$$H_n^*(x) = (-)^{[n/2]} \sum_{\nu=0}^{[n/2]} \frac{(-)^\nu}{4^\nu} \binom{n}{2\nu} \frac{(2\nu)!}{\nu!} x^{n-2\nu}$$

und daher gilt

$$x^n = (-)^{[n/2]} \sum_{\nu=0}^{[n/2]} \frac{(-)^\nu}{4^\nu} \binom{n}{2\nu} \frac{(2\nu)!}{\nu!} H_{n-2\nu}^*(x).$$

Um zu beweisen, daß die TSCHEBYSCHEFF*schen Polynome*

$$T_0(x) = 1, \qquad T_n(x) = \frac{\cos(n \cdot \arccos x)}{2^{n-1}} \qquad n \geq 1$$

kein involutorisches Polynomsystem bilden, benutzen wir das Korollar von Satz I, indem wir zeigen, daß für das Wertepaar (4, 0) die beiden Seiten in (5) verschiedenen Betrag haben. Das vierte TSCHEBYSCHEFFsche Polynom lautet: $T_4 = x^4 - x^2 + \frac{1}{8}$. Da

$$(x^4, T_0) = \int_{-1}^{+1} \frac{x^4 T_0}{\sqrt{1-x^2}} dx = \frac{3}{8}\pi, \qquad (T_0, T_0) = \pi$$

ist, folgt:

$$\frac{1}{8} = |p_0^{(4)}| \neq \frac{(x^4, T_0)}{(T_0, T_0)} = \frac{3}{8}.$$

Ähnlich schließen wir im Fall der LEGENDRE*schen Polynome*

$$P_n(x) = \frac{1}{\binom{2n}{n}} \sum_{\nu=0}^{[n/2]} (-)^\nu \binom{2n-2\nu}{n} \binom{n}{\nu} x^{n-2\nu}.$$

Das vierte Polynom lautet: $P_4 = x^4 - \frac{6}{7}x^2 + \frac{3}{35}$. Für (x^4, P_0) ergibt sich:

$$(x^4, P_0) = \int_{-1}^{+1} x^4 P_0 \, dx = \int_{-1}^{+1} x^4 \, dx = \frac{2}{5}.$$

Da $(P_0, P_0) = 2$ ist, folgt

$$\frac{3}{35} = |p_0^{(4)}| \neq \frac{(x^4, P_0)}{(P_0, P_0)} = \frac{1}{5},$$

d.h. auch die LEGENDREschen Polynome bilden kein involutorisches Polynomsystem. Hieraus können wir sofort schließen, daß auch die GEGENBAUERschen Polynome $C_n^{(\nu)}(x)$ für allgemeinen Parameterwert ν nicht involutorisch sind. Denn wäre dies der Fall, so insbesondere auch für $\nu = \frac{1}{2}$. $C_n^{(\frac{1}{2})}$ sind aber gerade die LEGENDREschen Polynome, für die wir oben das Gegenteil bewiesen haben.

Auch die JACOBI*schen Polynome*

$$\mathfrak{F}_n(\alpha, \gamma; x) = (-)^n \frac{\gamma \cdot \overline{\gamma+1} \dots \overline{\gamma+n-1}}{\overline{\alpha+n} \dots \overline{\alpha+2n-1}} F(-n, \alpha+n, \gamma; x)$$

bilden für allgemeine Parameterwerte α und β kein involutorisches System. Denn wählen wir etwa $\alpha = \gamma = 1$, so gilt wegen

$$(x^2, \mathfrak{F}_0(1,1;x)) = \int_0^1 x^2 \, dx = \tfrac{1}{3}$$

$$(\mathfrak{F}_0(1,1;x), \mathfrak{F}_0(1,1;x)) = 1$$

und $\mathfrak{F}_2(1,1;x) = x^2 - x + \frac{1}{6}$:

$$\frac{1}{6} = |p_0^{(2)}| \neq \left| \frac{(x^2, \mathfrak{F}_0)}{(\mathfrak{F}_0, \mathfrak{F}_0)} \right| = \frac{1}{3}.$$

Zusammenfassend können wir formulieren:

Satz II. *Unter den klassischen orthogonalen Polynomsystemen*

$$\{P_n(x)\}, \{\mathfrak{F}_n(\alpha, \gamma; x)\}, \{L_n^{(\alpha)}(x)\}, \{H_n\}, \{T_n\}, \{C_n^{(\nu)}\}$$

bilden das System der LAGUERRE*schen Polynome* $\{L_n^{(\alpha)}(x)\}$ *und das System der* HERMITE*schen Polynome* $\{H_n(x)\}$ *die einzigen involutorischen Polynomsysteme.*

§ 4. Das den Laguerreschen und Hermiteschen Polynomen übergeordnete Polynomsystem $\{P_{k,n}(x)\}$

Wie in der Einleitung erwähnt wurde, existiert ein von einem ganzzahligen Parameter $k \geq 1$ abhängiges orthogonales Polynomsystem $\{P_{k,n}(x)\}$,

das für $k=1$ und $k=2$ das LAGUERREsche ($\alpha=0$) bzw. HERMITEsche System enthält:

$$\{P_{1,n}(x)\}=\{L_n(x)\},\quad \{P_{2,n}(x)\}=\{H_n(x)\}.$$

Das System $\{P_{k,n}(x)\}$ ist mit der Belegungsfunktion $p(z)=e^{-z^k}$ orthogonal auf einem k-strahligen symmetrischen Integrationsstern der z-Ebene $\mathfrak{S}=\bigcup_{\lambda=0}^{k-1}\mathfrak{S}_\lambda$, wo $\mathfrak{S}_\lambda$ den Strahl $[z]=\left[e^{\frac{2\pi i}{k}\lambda}\cdot r\right]_{\leq r\leq\infty}$ bedeutet:

$$\left(P_{k,n}(z),P_{k,m}(z)\right)_{e^{-z^k}}^{(k)}=\sum_{\lambda=0}^{k-1}\int\limits_{\mathfrak{S}_\lambda}^{0\ \infty}P_{k,n}(z)\,\overline{P_{k,m}(z)}\,e^{-z^k}\,dr=0\qquad m\neq n.$$

Die Polynome $P_{k,n}(x)$ lauten, wenn wir $n=n_k^*+n_k k$ setzen $(0\leq n_k^*<k)$

$$P_{k,n}(x)=n_k!\cdot\sum_{\nu=0}^{n_k}\frac{(-)^{n_k-\nu}}{\nu!}\binom{n_k+\dfrac{2n_k^*+1}{k}-1}{n_k-\nu}x^{n_k^*+\nu k},\tag{6}$$

ihre Norm ist

$$(P_{k,n},P_{k,n})_{e^{-z^k}}^{(k)}=k\cdot(n_k!)^2\binom{n_k+\dfrac{2n_k^*+1}{k}-1}{n_k}\int_0^\infty r^{2n_k^*}e^{-r^k}\,dr.$$

Da die Polynomsysteme $\{P_{1,n}(x)\}$ und $\{P_{2,n}(x)\}$ involutorisch sind, liegt die Frage nahe, ob dies auch für alle k der Fall ist. Zur Anwendung des Kriteriums (5) berechnen wir das skalare Produkt:

$$\left(z^n,P_{k,\nu}(z)\right)=\sum_{\lambda=0}^{k-1}\int\limits_{\mathfrak{S}_\lambda}^{0\ \infty}z^n\,\overline{P_{k,\nu}(z)}\,e^{-r^k}\,dr$$

$$=\nu_k!\sum_{\lambda=0}^{k-1}\sum_{\varkappa=0}^{\nu_k}\frac{(-)^{\nu_k-\varkappa}}{\varkappa!}\binom{\nu_k+\dfrac{2\nu_k^*+1}{k}-1}{\nu_k-\varkappa}\int_0^\infty\left(re^{\frac{2\pi i}{k}\lambda}\right)^n\overline{\left(r\cdot e^{\frac{2\pi i}{k}\lambda}\right)^{\nu_k^*+\varkappa k}}\cdot e^{-r^k}\,dr$$

$$=\nu_k!\sum_{\varkappa=0}^{\nu_k}\frac{(-)^{\nu_k-\varkappa}}{\varkappa!}\binom{\nu_k+\dfrac{2\nu_k^*+1}{k}-1}{\nu_k-\varkappa}\cdot\int_0^\infty r^{n+\nu_k^*+\varkappa k}\left\{\sum_{\lambda=0}^{k-1}\left[e^{\frac{2\pi i}{k}(n-\nu_k^*-\varkappa k)}\right]^\lambda\right\}e^{-r^k}\,dr.$$

Da

$$\sum_{\lambda=0}^{k-1}\left[e^{\frac{2\pi i}{k}(n-\nu_k^*-\varkappa k)}\right]^\lambda=\begin{cases}\displaystyle\sum_{\lambda=0}^{k-1}1=k & \text{für } \nu\equiv n(k)\\[2ex] \dfrac{e^{\frac{2\pi i}{k}(n-\nu_k^*-\varkappa k)\cdot k}-1}{e^{\frac{2\pi i}{k}(n-\nu_k^*-\varkappa k)}-1}=0 & \text{für } \nu\not\equiv n(k)\end{cases}\tag{7}$$

ist $(z^n,P_{k,\nu})_{e^{-z^k}}^{(k)}\neq 0$ nur für $\nu\equiv n(k)$. Für diesen Fall gilt also

$$(z^n,P_{k,\nu})_{e^{-z^k}}^{(k)}=k\cdot\nu_k!\sum_{\varkappa=0}^{\nu_k}\frac{(-)^{\nu_k-\varkappa}}{\varkappa!}\binom{\nu_k+\dfrac{2\nu_k^*+1}{k}-1}{\nu_k-\varkappa}\int_0^\infty r^{n+\nu_k^*+\varkappa k}e^{-r^k}\,dr.$$

Setzen wir $\Gamma_k(z)=\int_0^\infty r^{z-1}e^{-r^k}dr$, so erhalten wir durch Rückführung auf die Γ-Funktion ($k=1$) die Funktionalgleichung:

$$\Gamma_k(z)=\frac{\overline{z-k}\cdot\overline{z-2k}\dots\overline{z-\lambda k}}{k^\lambda}\cdot\Gamma_k(z-\lambda k)$$

(für $k=1$ liegt die Funktionalgleichung der Γ-Funktion vor) und damit für

$$\int_0^\infty r^{2n_k^*+(n_k+\varkappa)k}e^{-r^k}dr=\frac{\overline{2n_k^*+1+\overline{n_k+\varkappa-1}\cdot k}\dots\overline{2n_k^*+1}}{k^{n_k+\varkappa}}\int_0^\infty r^{2n_k^*}e^{-r^k}dr$$

$$=\binom{n_k+\varkappa+\frac{2n_k^*+1}{k}-1}{n_k+\varkappa}(n_k+\varkappa)!\int_0^\infty r^{2n_k^*}e^{-r^k}dr.$$

Es wird also

$$\left(z^n,P_{k,\nu}\right)_{e^{-z^k}}^{(k)}=\nu_k!\cdot k\cdot\int_0^\infty r^{2n_k^*}e^{-r^k}dr\times$$

$$\times\sum_{\varkappa=0}^{\nu_k}\frac{(-)^{\nu_k-\varkappa}}{\varkappa!}\binom{\nu_k+\frac{2n_k^*+1}{k}-1}{\nu_k-\varkappa}(n_k+\varkappa)!\binom{n_k+\varkappa+\frac{2n_k^*+1}{k}-1}{n_k+\varkappa}$$

$$=\nu_k!\cdot k\cdot\int_0^\infty r^{2n_k^*}e^{-r^k}dr\cdot(-)^{\nu_k}\cdot n_k!\binom{\nu_k+\frac{2\nu_k^*+1}{k}-1}{\nu_k}\binom{n_k+\frac{2n_k^*+1}{k}-1}{n_k}\times$$

$$\times F\left(-\nu_k,\ n_k+\frac{2n_k^*+1}{k},\ \frac{2n_k^*+1}{k};1\right)$$

$$=k\cdot\nu_k!\int_0^\infty r^{2n_k^*}e^{-r^k}dr\binom{n_k}{\nu_k}\frac{\Gamma\left(n_k+\frac{2n_k^*+1}{k}\right)}{\Gamma\left(\frac{2n_k^*+1}{k}\right)}.$$

Nun ist nach (6) $p_\nu^{(n)}=0$ für $\nu\not\equiv n(k)$ und damit nach (7) die Beziehung (5) für solche ν immer erfüllt. Ist $\nu\equiv n(k)$, so lautet (5) für n und $\nu=n_k^*+\varkappa k$ ($n=0,1,2,\dots;\ \varkappa=0,1,\dots,n_k$), da

$$\left(P_{k,n_k^*+\varkappa k},P_{k,n_k^*+\varkappa k}\right)_{e^{-z^k}}^{(k)}=k(\varkappa!)^2\binom{\varkappa+\frac{2n_k^*+1}{k}-1}{\varkappa}\int_0^\infty r^{2n_k^*}e^{-r^k}dr$$

ist:

$$C_n\cdot C_{n_k^*+\varkappa k}\cdot n_k!\frac{(-)^{n_k-\varkappa}}{\varkappa!}\binom{n_k+\frac{2n_k^*+1}{k}-1}{n_k-\varkappa}=\frac{n_k!\binom{n_k+\frac{2n_k^*+1}{k}-1}{n_k}\binom{n_k}{\varkappa}}{\varkappa!\binom{\varkappa+\frac{2n_k^*+1}{k}-1}{\varkappa}}$$

oder:

(8) $$C_n\cdot C_{n_k^*+\varkappa k}=(-)^{n_k}\cdot(-)^\varkappa\qquad(n=0,1,\dots;\ \varkappa=0,1,\dots,n_k).$$

Diese Gleichung kann durch verschiedene Vorzeichensysteme $\{c_n\}$ gelöst werden. Setzen wir $\varkappa = 0$, so folgt $c_n = (-)^{n_k} v_{n_k^*}$. Wie wir auch bei k Vorzeichen $\{c_0, c_1, \ldots, c_{k-1}\}^n$ wählen, stets ist (8) durch $c_n = (-)^{n_k} v_{n_k^*}$ $(n = 0, 1, \ldots)$ gelöst, und alle Lösungen lassen sich auch so darstellen. Es gilt also: für alle k $(k = 1, 2, \ldots)$:

Das Polynomsystem $\{P_{k,n}(x)\}$ ist involutorisch. Es besitzt genau 2^k involutorische Formen.

Das einfachste System ergibt sich mit $c_0 = c_1 = \cdots = c_{k-1} = 1$.

$$P^*_{k,n}(x) = n_k! \sum_{\nu=0}^{n_k} \frac{(-)^\nu}{\nu!} \binom{n_k + \dfrac{2n_k^* + 1}{k} - 1}{n_k - \nu} x^{n_k^* + \nu k}.$$

Für $k = 1$ und $k = 2$ erhält man hieraus gerade die in § 1 bzw. § 3 angegebenen involutorischen Formen der LAGUERREschen bzw. HERMITEschen Polynome.

Gießen, Mathematisches Institut der Universität

(Eingegangen am 15. September 1955)

SCHOENEBERG, B.
Math. Zeitschr. Bd. 65, S. 16—24 (1956)

Über die Diskriminante der elliptischen Funktionen und ihre Quadratwurzel

Von

BRUNO SCHOENEBERG

Die Fourier-Koeffizienten von

$$\Delta(\tau) = e^{2\pi i\tau} \prod_{n=1}^{\infty} (1 - e^{2\pi i\tau n})^{24} = \sum_{n=1}^{\infty} \tau(n)\, e^{2\pi i\tau n}, \quad \Im\mathrm{m}\ \tau > 0$$

sind ihrer arithmetischen Natur nach noch nicht in wünschenswertem Umfange bekannt. Man weiß durch MORDELL[1]), daß zwischen ihnen folgende Beziehungen bestehen:

(1) $\tau(m\,n) = \tau(m)\,\tau(n)$ bei $(m, n) = 1$,

(2) $\tau(p^{\alpha}) = \tau(p)\,\tau(p^{\alpha-1}) - p^{11}\tau(p^{\alpha-2})$ bei primem p und ganzem $\alpha \geq 2$.

Dieses Ergebnis hat HECKE[2]) erneut hergeleitet im Rahmen einer allgemeineren Theorie, der man unter anderem entnimmt, daß eine (1) und (2) entsprechende Beziehung auch für

$$\sqrt{\Delta(\tau)} = e^{\frac{2\pi i\tau}{2}} \prod_{n=1}^{\infty} (1 - e^{2\pi i\tau n})^{12} = \sum_{\substack{n=1 \\ n \equiv 1\,(2)}}^{\infty} \tau_2(n)\, e^{\frac{2\pi i\tau}{2} n}, \quad \Im\mathrm{m}\ \tau > 0$$

gilt. Es ist nämlich

(3) $\tau_2(m\,n) = \tau_2(m)\,\tau_2(n)$ bei $(m, n) = 1$,

(4) $\tau_2(p^{\alpha}) = \tau_2(p)\,\tau_2(p^{\alpha-1}) - p^5\,\tau_2(p^{\alpha-2})$ bei primem $p \neq 2$ und $\alpha \geq 2$.

Durch (1) und (3) ist die Frage nach dem Verschwinden von $\tau(n)$ und $\tau_2(n)$ auf die Frage nach dem Verschwinden von $\tau(p^{\alpha})$ und $\tau_2(p^{\alpha})$ zurückgeführt. Nun ist nach MORDELL[1])

(5) $$\tau(p^{\alpha}) = \frac{p^{\frac{11}{2}\alpha} \sin(\alpha + 1)\,\Theta_p}{\sin\Theta_p} \quad \text{mit} \quad 2\cos\Theta_p = \frac{\tau(p)}{p^{\frac{11}{2}}}, \quad (\alpha \geq 0),$$

---

[1]) MORDELL, L. J.: On Mr. Ramanujan's empirical expansions of modular functions. Proc. Cambridge Philos. Soc. **19**, 117—124 (1917).

[2]) HECKE, E.: Über Modulfunktionen und die Dirichletschen Reihen mit Eulerscher Produktentwicklung. I, II. Math. Ann. **114**, 1—28, 316—351 (1937).

wie man sofort bestätigt, wenn man diesen Ausdruck für $\tau(p^2)$ in (2) einsetzt. Entsprechend folgt aus (4)

$$(6) \qquad \tau_2(p^2) = \frac{p^{\frac{5}{2}\alpha} \sin(\alpha+1)\Theta'_p}{\sin\Theta'_p} \qquad \text{mit} \qquad 2\cos\Theta'_p = \frac{\tau_2(p)}{p^{\frac{5}{2}}}.$$

Aus (5) folgt, wie LEHMER[3]) gezeigt hat, daß, wenn $\tau(n)$ überhaupt einmal verschwindet, das kleinste derartige n eine Primzahl ist. Bei verschwindendem $\tau(p^2)$ ist nämlich Θ_p reell, also $\left|\frac{\tau(p)}{p^{\frac{11}{2}}}\right| \leq 2$. Andererseits ist bei verschwindendem $\tau(p^2)$, wie sich aus (2) ergibt, $\frac{\tau(p)}{p^{\frac{11}{2}}}$ ganz-algebraisch. Also ist in $\left(\frac{\tau(p)}{p^{\frac{11}{2}}}\right)^2 - r = 0$ das rationale r eine ganze Zahl. Überdies ist $0 \leq r \leq 4$. Hier sind $r = 1$ und 4 nicht möglich, weil $p^{\frac{11}{2}}$ nicht rational ist, und $r = 2$ und 3 nicht, weil $\tau(2) \neq \pm 2^6$ und $\tau(3) \neq \pm 3^6$ ist. Also ist $\tau(p^2) = 0$ bei $\tau(p) \neq 0$ nicht möglich. Auf völlig entsprechende Weise ergibt sich: Das kleinste $n \equiv 1(2)$ mit $\tau_2(n) = 0$ ist eine Primzahl. Ob es ein verschwindendes $\tau(n)$ oder ein verschwindendes $\tau_2(n)$ mit $n \equiv 1(2)$ gibt, weiß man nicht. Für $\tau(n)$ ist eine Reihe von Kongruenzeigenschaften bekannt, aus denen für das Verschwinden von $\tau(n)$ einschränkende Bedingungen folgen. In der vorliegenden Arbeit wird eine weitere Kongruenzeigenschaft hergeleitet. Da bei $\tau_2(n)$ ähnliche Probleme wie bei $\tau(n)$ vorliegen, soll hier auch die Funktion $\sqrt{\Delta(\tau)}$ nach derselben Methode untersucht werden.

Die Fourier-Koeffizienten von $\Delta(\tau)$ und $\sqrt{\Delta(\tau)}$ werden hier mit Hilfe gewisser positiv-definiter quadratischer Formen in 8 und 4 Variablen unter Heranziehung von Kugelfunktionen gewonnen. Dazu sei an folgenden Sachverhalt erinnert[4]):

Es sei $\mathfrak{A} = (a_{ik})$ eine ganzzahlige symmetrische Matrix mit $2r$ Zeilen und Spalten und $a_{ii} \equiv 0(2)$; die quadratische Form $\sum_{i,k=1}^{2r} a_{ik} x_i x_k$ sei positiv-definit; die Determinante von $\mathfrak{A}$ sei $|a_{ik}| = D$ und $(a_{ik})^{-1} = \frac{1}{D}(A_{ik})$. Der größte gemeinsame Teiler $\left(\frac{A_{11}}{2}, \ldots, \frac{A_{2r,2r}}{2}, \ldots, A_{ik} \ldots\right)$ sei K, und es sei $N = D/K$. Ferner sei $P_\nu(\mathfrak{x})$ eine Kugelfunktion in $\mathfrak{x} = (x_1, \ldots, x_{2r})$ in bezug auf $\sum_{i,k=1}^{2r} a_{ik} x_i x$ von der Dimension ν. Damit ist gemeint, daß $P_\nu(\mathfrak{x})$ durch eine Transformation, welche $\sum a_{ik} x_i x_k$ in eine Summe von Quadraten transformiert, zu einer gewöhnlichen Kugelfunktion in den neuen Variablen wird. Weiter sei $\mathfrak{h}$ ein ganz-

[3]) LEHMER, D. H.: The vanishing of Ramanujan's function $\tau(n)$. Duke Math. J. **14**, 429—433 (1947).

[4]) SCHOENEBERG, B.: Das Verhalten von mehrfachen Thetareihen bei Modulsubstitutionen. Math. Ann. **116**, 511—523 (1939).

zahliger $2r$-dimensionaler Vektor mit $\mathfrak{A}\mathfrak{h} \equiv 0(N)$. Dann ist mit $(\mathfrak{n} = n_1, \ldots, n_{2r})$

$$(7) \qquad \begin{cases} \vartheta(\tau; \mathfrak{h}, \mathfrak{A}, P_\nu) = \sum\limits_{\mathfrak{n} \equiv \mathfrak{h}(N)} P_\nu(\mathfrak{n})\, e^{2\pi i \tau \frac{\frac{1}{2}\sum a_{ik} n_i n_k}{N^2}} = \sum\limits_{n=0}^{\infty} c_{P_\nu}(n)\, e^{\frac{2\pi i \tau}{N} n}, \\ c_{P_\nu}(n) = \sum\nolimits^{*} P_\nu(\mathfrak{n}), \end{cases}$$

wo für $c_{P_\nu}(n)$ über alle $\mathfrak{n} \equiv \mathfrak{h}(N)$ mit $\frac{1}{2N} \sum\limits_{i,k=1}^{2r} a_{ik} n_i n_k = n$ zu summieren ist, eine ganze Modulform der Dimension $-(r+\nu)$ zur Stufe N.

Daß für die Kugelfunktionen $(\sum a_{ik} l_i x_k)^\nu$ mit $\sum a_{ik} l_i l_k = 0$ auf diese Weise Modulformen entstehen, folgt aus dem Ansatz. Es führen aber alle Kugelfunktionen in bezug auf $\sum a_{ik} x_i x_k$ zu Modulformen, weil jede Kugelfunktion der Dimension ν linear homogen durch die $(\sum a_{ik} l_i x_k)^\nu$ mit der angegebenen Bedingung für die l_i ausdrückbar ist.

Für das Folgende brauchbare spezielle Kugelfunktionen in bezug auf $\sum a_{ik} x_i x_k = 2\boldsymbol{Q}(\mathfrak{x})$ gewinnt man durch den Ansatz

$$P_\nu(\mathfrak{x}) = \sum_{j=0}^{[\nu/2]} a_j x_1^{\nu-2j} \boldsymbol{Q}(\mathfrak{x})^j, \qquad \sum_{i=1}^{2r} \frac{\partial^2}{\partial y_i^2} P_\nu(\mathfrak{x}) = 0.$$

Dabei ist $x_i = \sum\limits_{k=1}^{2r} t_{ik} y_k$ eine Transformation, durch die $\sum\limits_{i,k=1}^{2r} a_{ik} x_i x_k = \sum\limits_{i=1}^{2r} y_i^2$ wird. Man erhält eine Rekursionsformel für die Koeffizienten a_j und damit

$$(8) \qquad P_\nu(\mathfrak{x}) = \sum_{j=0}^{[\nu/2]} \frac{(-1)^j}{j!} \left(\frac{A_{11}}{2D}\right)^j \frac{\nu(\nu-1)\ldots(\nu-2j+1)}{(r+\nu-2)\ldots(r+\nu-j-1)} \boldsymbol{Q}(\mathfrak{x})^j.$$

Um eine Entwicklung für $\Delta(\tau)$ zu erhalten, setzt man in (7) für $\mathfrak{A}$ die bekannte Matrix

$$\mathfrak{A}_8 = \begin{bmatrix} 2, & 1, & 0, & . & . & . & . & 0 \\ 1, & 4, & 1, & & & & & . \\ 0, & 1, & 12, & 3, & & & & . \\ 0, & 0, & 3, & 20, & 5, & & & . \\ 0, & & & 5, & 4, & 3, & & . \\ 0, & & & & 3, & 4, & 1, & . \\ 0, & & & & & 1, & 2, & 1 \\ 0, & . & . & . & . & . & 1, & 2 \end{bmatrix} \quad \text{mit} \quad D = K = N = 1$$

und für $P_\nu(\mathfrak{x})$ eine Kugelfunktion der Dimension 8 in bezug auf $\mathfrak{A}_8$. Dann ist die Funktion (7) eine ganze Modulform der Dimension -12 zur Stufe 1, die in der Spitze des Fundamentalbereichs verschwindet und daher, wenn sie nicht identisch verschwindet, bis auf einen konstanten Faktor gleich $\Delta(\tau)$ ist.

Die mit $\mathfrak{A}_8$ gebildeten Reihen (7) verschwinden für $\nu < 8$ identisch, und zwar für ungerades ν trivialerweise und für gerades ν als Spitzenformen einer Dimension > -12 zur Stufe 1. Es ist vorteilhaft, auch ihre Reihenentwicklungen heranzuziehen.

Die ersten vier Kugelfunktionen in bezug auf $\mathfrak{A}_8$ von gerader Dimension sind nach (8) mit $A_{11} = 2$

$$(9)\quad \begin{cases} P_2(\mathfrak{x}) = x_1^2 - \frac{1}{2}\,\boldsymbol{Q}_8(\mathfrak{x}) \\ P_4(\mathfrak{x}) = x_1^4 - 2\,x_1^2\,\boldsymbol{Q}_8(\mathfrak{x}) + \frac{2}{5}\,\boldsymbol{Q}_8(\mathfrak{x})^2 \\ P_6(\mathfrak{x}) = x_1^6 - \frac{15}{4}\,x_1^4\,\boldsymbol{Q}_8(\mathfrak{x}) + \frac{45}{14}\,x_1^2\,\boldsymbol{Q}_8(\mathfrak{x})^2 - \frac{5}{14}\,\boldsymbol{Q}_8(\mathfrak{x})^4 \\ P_8(\mathfrak{x}) = x_1^8 - \frac{28}{5}\,x_1^6\,\boldsymbol{Q}_8(\mathfrak{x}) + \frac{28}{3}\,x_1^4\,\boldsymbol{Q}_8(\mathfrak{x})^2 - \frac{14}{3}\,x_1^2\,\boldsymbol{Q}_8(\mathfrak{x})^3 + \frac{1}{3}\,\boldsymbol{Q}_8(\mathfrak{x})^4. \end{cases}$$

Die Bedingung $\mathfrak{A}\mathfrak{h} \equiv 0\,(N)$ entfällt hier. Die mit $\nu \equiv 0$ gebildete Reihe ist

$$\sum_{\mathfrak{n}} e^{2\pi i \tau \boldsymbol{Q}_s(\mathfrak{n})} = \sum_{n=0}^{\infty} z_8(n)\, e^{2\pi i \tau n},$$

wo $z_8(n)$ gleich der Anzahl der ganzzahligen Lösungen $\mathfrak{n}$ von $\boldsymbol{Q}_8(\mathfrak{n}) = n$ ist. Diese Reihe ist bis auf einen konstanten Faktor gleich der Eisensteinschen Reihe der Dimension -4. Das ergibt wegen $z_8(1) = 240$

$$(10)\qquad z_8(n) = 240\,\sigma_3(n),$$

wo, wie üblich, $\sigma_k(n) = \sum_{\substack{d>0 \\ d \mid n}} d^k$ ist. Benutzt man diesen Wert für $z_8(n)$, so folgen aus dem identischen Verschwinden der mit P_2, P_4, P_6 gebildeten Reihen nacheinander folgende Relationen:

$$(11)\quad \begin{cases} \sum\limits_{\boldsymbol{Q}_s(\mathfrak{n}) = n} n_1^2 = 120\, n\, \sigma_3(n), \\ \sum\limits_{\boldsymbol{Q}_s(\mathfrak{n}) = n} n_1^4 = 144\, n^2\, \sigma_3(n), \\ \sum\limits_{\boldsymbol{Q}_s(\mathfrak{n}) = n} n_1^6 = 240\, n^3\, \sigma_3(n). \end{cases}$$

Zur Berechnung jeder dieser Relationen werden die vorangehenden benutzt. Aus ihnen folgt mit P_8 aus (9)

$$\sum_{\mathfrak{n}} P_8(\mathfrak{n})\, e^{2\pi i \tau \boldsymbol{Q}_s(\mathfrak{n})} = \sum_{n=1}^{\infty} c_n\, e^{2\pi i \tau n}, \qquad c_n = \sum_{\boldsymbol{Q}_s(\mathfrak{n}) = n} n_1^8 - 480\, n^4\, \sigma_3(n).$$

Hier ist $c_1 = 144$. Denn aus $\sum\limits_{\boldsymbol{Q}_s = 1} n_1^2 = 120$ und $\sum\limits_{\boldsymbol{Q}_s = 1} n_1^4 = 144$ folgt, daß es 112 Lösungen von $\boldsymbol{Q}_1^8(\mathfrak{n}) = 1$ mit $n_1 = 1$ und 2 Lösungen mit $n_2 = 2$ gibt. Also ist $\sum\limits_{\boldsymbol{Q}_s = 1} n^2 = 624$. Damit ist[5])

$$(12)\qquad \tau(n) = \frac{1}{144}\left\{ \sum_{\boldsymbol{Q}_s(\mathfrak{n}) = n} n_1^8 - 480\, n^4\, \sigma_3(n) \right\}.$$

[5]) Vgl. hierzu HECKE, E.: Analytische Theorie der positiven quadratischen Formen. Kgl. Danske Vidensk. Selsk. Math.-phys. Medd. **17**, 12 (1940), bes. S. 86—89.

Eine weitere Darstellung von $\Delta(\tau)$ gewinnt man mit Hilfe der quadratischen Form in 16 Variablen

$$Q_{16}(\mathfrak{x}) = Q_8(\mathfrak{x}_1) + Q_8(\mathfrak{x}_2)$$

mit dem schon eingeführten Q_8 und $\mathfrak{x}_1 = (x_1, \ldots, x_8)$, $\mathfrak{x}_2 = (x_9, \ldots, x_{16})$, $\mathfrak{x} = (x_1, \ldots, x_{16})$. Für diese Form ist

$$\mathfrak{A}_{16} = \begin{pmatrix} \mathfrak{A}_8 & 0 \\ 0 & \mathfrak{A}_8 \end{pmatrix}, \qquad D = K = N = 1.$$

Die ersten beiden Kugelfunktionen gerader Dimension sind

$$\begin{cases} P_2'(\mathfrak{x}) = x_1^2 - \frac{1}{4} Q_{16}(\mathfrak{x}), \\ P_4'(\mathfrak{x}) = x_1^4 - \frac{6}{5} x_1^2 Q_{16}(\mathfrak{x}) + \frac{9}{175} Q_{16}(\mathfrak{x})^2. \end{cases} \tag{13}$$

Die mit $\nu = 0$ gebildete Reihe (7) ist bis auf einen konstanten Faktor gleich der Eisensteinschen Reihe der Dimension -8. Das ergibt für die Anzahl $z_{16}(n)$ der Lösungen von $Q_{16}(\mathfrak{n}) = n$

$$z_{16}(n) = 480\, \sigma_7(n). \tag{14}$$

Die mit $P_2'(\mathfrak{x})$ gebildete Reihe verschwindet als Spitzenform der Dimension -10 identisch. Also ist

$$\sum_{Q_{16}(\mathfrak{n}) = n} n_1^2 = 120\, n\, \sigma_7(n). \tag{15}$$

Damit folgt für die Koeffizienten c_n' der mit $P_4'(\mathfrak{x})$ gebildeten Reihe

$$\sum_{\mathfrak{n}} P_4'(\mathfrak{n})\, e^{2\pi i \tau Q_{16}(\mathfrak{n})} = \sum_{n=1}^{\infty} c_n' e^{2\pi i \tau n}, \qquad c_n' = \sum_{Q_{16}(\mathfrak{n}) = n} n_1^4 - 80\, n^2 \sigma_7(n).$$

Hier ist $c_1' = 64$, also ist die Reihe eine nicht identisch verschwindende Spitzenform der Dimension -12 und damit

$$\tau(n) = \frac{1}{64} \left\{ \sum_{Q_{16}(\mathfrak{n}) = n} n_1^4 - 80\, n^2 \sigma_7(n) \right\}. \tag{16}$$

Zur weiteren Untersuchung von $\tau(n)$ werde $\sum_{Q_{16}(\mathfrak{n}) = n} n_1^4$ umgeformt. Es ist

$$\begin{aligned} \sum_{Q_{16}(\mathfrak{n}) = n} n_1^4 &= \sum_{Q_8(\mathfrak{n}_1) = n} n_1^4 + \sum_{Q_8(\mathfrak{n}_1) = n-1} n_1^4 \cdot z_8(1) + \sum_{Q_8(\mathfrak{n}_1) = n-2} n_1^4 \cdot z_8(2) + \cdots \\ &= \sum_{Q_8(\mathfrak{n}_1) = n} n_1^4 + \sum_{i=1}^{n-1} \sum_{Q_8(\mathfrak{n}_1) = n-i} n_1^4\, z_8(i). \end{aligned}$$

Damit ist nach (11) und (10)

$$\sum_{Q_{16}(\mathfrak{n}) = n} n_1^4 = 144\, n^2 \sigma_3(n) + 144 \cdot 240 \sum_{i=1}^{n-1} (n-i)^2 \sigma_3(n-i)\, \sigma_3(i)$$

und nach (16)

$$\tau(n) = \frac{n^2}{4}\{9\sigma_3(n) - 5\sigma_7(n)\} + 9\cdot 240 \sum_{i=1}^{n-1} (n-i)^2 \sigma_3(n-i)\,\sigma_3(i). \tag{17}$$

Das ergibt für $\tau(n)$ die *Kongruenzeigenschaft*

$$\tau(n) \equiv \frac{n^2}{4}\{9\sigma_3(n) - 5\sigma_7(n)\} \bmod 9\cdot 240. \tag{18}$$

Für eine Primzahl $p \neq 2, 3, 5$ ist also $\tau(p)$ höchstens dann gleich Null, wenn die Kongruenz

$$5p^7 - 9p^3 - 4 \equiv 0 \bmod 2^6 \cdot 3^3 \cdot 5$$

erfüllt ist, d.h. höchstens für

$$p \equiv -1 \bmod 2^2 \cdot 3^3 \cdot 5. \tag{19}$$

Das ist eine Kongruenzeigenschaft von bekanntem Typus. Die Methode ihrer Herleitung ist auch auf die Funktion $\sqrt{\Delta(\tau)}$ anwendbar.

$\sqrt{\Delta(\tau)}$ ist eine ganze Modulform der Dimension -6 zur Stufe 2, die in allen Spitzen des Fundamentalbereiches verschwindet, und dadurch bis auf einen konstanten Faktor bestimmt. Zu ihrer Darstellung wird die quadratische Form $Q_4(\mathfrak{x})$ mit

$$\mathfrak{A}_4 = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix}, \quad D = 4,\ K = N = 2$$

herangezogen. Die Kugelfunktionen bezüglich dieser Form von der zweiten und vierten Dimension sind

$$\begin{cases} R_2(\mathfrak{x}) = x_1^2 - \frac{1}{2} Q_4(\mathfrak{x}), \\ R_4(\mathfrak{x}) = x_1^4 - \frac{3}{2} x_1^2 Q_4(\mathfrak{x}) + \frac{1}{4} Q_4(\mathfrak{x})^2. \end{cases} \tag{20}$$

Eine brauchbare Lösung von $\mathfrak{A}_4 \mathfrak{h} \equiv 0\,(2)$ ist $\mathfrak{h} \equiv (1, 1, 0, 0) \bmod 2$. Die mit $\nu = 0$ gebildete Reihe ist

$$\sum_{\mathfrak{n} \equiv \mathfrak{h} \bmod 2} e^{\frac{2\pi i \tau}{2} \frac{Q_4(\mathfrak{n})}{2}} = \sum_{n=0}^{\infty} z_4(n)\, e^{\frac{2\pi i \tau}{2} n},$$

in der nur Exponenten $n \equiv 1 \bmod 2$ auftreten. Die Anzahl $z_4(n)$ der Lösungen von $\frac{1}{2} Q_4(\mathfrak{n}) = n$ mit $\mathfrak{n} \equiv \mathfrak{h} \bmod 2$ bestimmt man mit Hilfe der p-Teilwerte zur Stufe 2 zu

$$z_4(n) = 8\sigma_1(n), \quad n \equiv 1 \bmod 2. \tag{21}$$

Die Reihe mit $R_2(\mathfrak{x})$ ist als Spitzenform der Dimension -4 zur Stufe 2 identisch Null. Damit ist bei im folgenden immer als ungerade vorausgesetztem n.

$$\sum_{\substack{\frac{1}{2}Q_4(\mathfrak{n})=n \\ \mathfrak{n}\equiv\mathfrak{h} \bmod 2}} n_1^2 = 8n\,\sigma_1(n). \tag{22}$$

Für die Koeffizienten d_n der mit $R_4(\mathfrak{x})$ gebildeten Reihe folgt bei Benutzung von (21) und (22)

$$\sum_{\mathfrak{n}\equiv\mathfrak{h} \bmod 2} R_4(\mathfrak{n})\, e^{\frac{2\pi i\tau}{2}\frac{Q_4(\mathfrak{n})}{2}} = \sum_{n=1}^{\infty} d_n\, e^{\frac{2\pi i\tau}{2}n}, \qquad d_n = \sum_{\substack{\frac{1}{2}Q_4(\mathfrak{n})=n \\ \mathfrak{n}\equiv\mathfrak{h} \bmod 2}} n_1^4 - 16n^2\sigma_1(n).$$

Hier ist $d_1 = -8$, wie aus (22) folgt, wenn man dort $n=1$ setzt. Damit ist

$$\tau_2(n) = -\frac{1}{8}\left\{\sum_{\substack{\frac{1}{2}Q_4(\mathfrak{n})=n \\ \mathfrak{n}\equiv\mathfrak{h} \bmod 2}} n_1^4 - 16n^2\sigma_1(n)\right\}. \tag{23}$$

Die Summe besteht aus $8\sigma_1(n)$ Summanden, die alle kongruent 1 mod 16 sind. Für den Fall, daß $\tau_2(n) = 0$ ist, muß also $8\sigma_1(n) \equiv 0 \bmod 16$ sein. Daraus folgt: $\tau_2(n)$ *ist sicher dann von Null verschieden, wenn n eine ungerade Anzahl von Teilern hat, wenn also n ein Quadrat ist.*

Eine weitere Darstellung für $\sqrt{\Delta(\tau)}$ gewinnt man durch die quadratische Form in 8 Variablen $Q'_8(\mathfrak{x}) = Q_4(\mathfrak{x}_1) + Q_4(\mathfrak{x}_2)$ mit

$$\mathfrak{A}'_8 = \begin{bmatrix} \mathfrak{A}_4 & 0 \\ 0 & \mathfrak{A}_4 \end{bmatrix}, \qquad D = 16,\ K = 8,\ N = 2.$$

Es werde $\mathfrak{h} \equiv (1, 1, 0, 0, \ldots, 0) \bmod 2$ gesetzt. Die Kugelfunktion der Dimension 2 ist jetzt

$$R'_2(\mathfrak{x}) = x_1^2 - \tfrac{1}{4}\,Q'_8(\mathfrak{x}). \tag{24}$$

Die Anzahl $z'_8(n)$ der Lösungen von $\frac{1}{2}Q'_8(\mathfrak{n}) = n$ mit $\mathfrak{n} \equiv \mathfrak{h} \bmod 2$ ist gleich

$$z'_8(n) = 8\sigma_3(n), \tag{25}$$

wie man sieht, wenn man die Eisensteinschen Reihen der Dimension -4 zur Stufe 2 heranzieht. Mit (24) und (25) ist

$$\sum_{\mathfrak{n}\equiv\mathfrak{h} \bmod 2} R'_2(\mathfrak{n})\, e^{\frac{2\pi i\tau}{2}\frac{Q'_4(\mathfrak{n})}{2}} = \sum_{n=1}^{\infty} d'_n\, e^{\frac{2\pi i\tau}{2}n}, \qquad d'_n = \sum_{\substack{\frac{1}{2}Q'_4(\mathfrak{n})=n \\ \mathfrak{n}\equiv\mathfrak{h} \bmod 2}} n_1^2 - 4n\,\sigma_3(n).$$

Wegen $d'_1 = 4$ ist also

$$\tau_2(n) = \frac{1}{4}\left\{\sum_{\substack{\frac{1}{2}Q'_4(\mathfrak{n})=n \\ \mathfrak{n}\equiv\mathfrak{h} \bmod 2}} n_1^2 - 4n\,\sigma_3(n)\right\}. \tag{26}$$

Die hier auftretende Summe läßt sich umformen zu

$$\sum_{\substack{n \\ \frac{1}{2}Q_4(n)=n \\ n\equiv 0 \bmod 2}} n_1^2 = \sum_{\substack{n_1 \\ \frac{1}{2}Q_4(n_1)=n \\ n_1\equiv(1\,1\,0\,0)\bmod 2}} n_1^2 + \sum_{\substack{n_1 \\ \frac{1}{2}Q_4(n_1)=n-2 \\ n_1\equiv(1\,1\,0\,0)\bmod 2}} n_1^2 \cdot z_4'(2) + \sum_{\substack{n_1 \\ \frac{1}{2}Q_4(n_1)=n-4 \\ n_1\equiv(1\,1\,0\,0)\bmod 2}} n_1^2 \cdot z_4'(4) + \cdots.$$

Dabei ist $z_4'(2i)$ die Anzahl der Lösungen von $\frac{1}{2}Q_4(n_2) = 2i$ mit $n_2 \equiv (0\,0\,0\,0) \bmod 2$, also gleich $24\sigma_1'(i)$ mit $\sigma_1'(i) = \sum_{\substack{d\mid i,\ d>0 \\ d\equiv 1 \bmod 2}} d$. Hieraus und aus (22) folgt

$$\sum_{\substack{n \\ \frac{1}{2}Q_4(n)=n \\ n\equiv 0 \bmod 2}} n_1^2 = 8n\,\sigma_1(n) + \sum_{i=1}^{\frac{n-1}{2}} 8(n-2i)\,\sigma_1(n-2i)\cdot 24\,\sigma_1'(i)$$

und schließlich

$$\tau_2(n) = n\left(2\sigma_1(n) - \sigma_3(n)\right) + 48\sum_{i=1}^{\frac{n-1}{2}} (n-2i)\,\sigma_1(n-2i)\,\sigma_1'(i). \tag{27}$$

Die Koeffizienten $\tau_2(n)$ besitzen demnach die *Kongruenzeigenschaft*

$$\tau_2(n) \equiv n\left(2\sigma_1(n) - \sigma_3(n)\right) \bmod 48. \tag{28}$$

Wenn für primzahliges $p \neq 2, 3$ der Koeffizient $\tau_2(p) = 0$ ist, besteht die Kongruenz

$$p^3 - 2p^2 - 1 \equiv 0 \bmod 48,$$

die nur für $p \equiv -1 \bmod 48$ erfüllt ist.

Eine weitere Darstellung von $\sqrt{\Delta(\tau)}$ gewinnt man durch die quadratische Form in 12 Variablen $Q_{12}(\mathfrak{x}) = Q_4(\mathfrak{x}_1) + Q_4(\mathfrak{x}_2) + Q_4(\mathfrak{x}_3)$, wenn man die Eisensteinsche Reihe $G_6(\tau) = \sum_{\substack{n=1 \\ n\equiv 1 \bmod 2}}^{\infty} \sigma_5(n)\, e^{\frac{2\pi i\tau}{2}n}$ der Dimension -6 zur Stufe 2 heranzieht. Durch Berechnung der ersten Koeffizienten erhält man

$$\sqrt{\Delta(\tau)} = G_6(\tau) - \frac{1}{2}\sum_{\substack{n \\ n_i\equiv(1\,1\,0\,0)\bmod 2}} e^{\frac{2\pi i\tau}{2}\cdot\frac{1}{2}Q_{12}(n)}.$$

Diese Darstellung führt zu

$$\tau_2(n) = \sigma_5(n) - \tfrac{1}{2}z_{12}(n), \tag{29}$$

wo $z_{12}(n)$ die Anzahl der Lösungen von $\frac{1}{2}Q_{12}(n) = n$ mit $n_i \equiv (1\,1\,0\,0) \bmod 2$ ist. Wegen $z_{12}(n) \equiv 0 \bmod 8^3$, wie aus (21) folgt, ist

$$\tau_2(n) \equiv \sigma_5(n) \bmod 2^8. \tag{30}$$

Es kann danach $\tau_2(p)$ höchstens dann verschwinden, wenn $p \equiv -1 \bmod 2^8$ ist.

Schließlich ist noch eine Darstellung von $\sqrt{\Delta(\tau)}$ gegeben durch

$$\sqrt{\Delta(\tau)} = \frac{1}{8}\left\{ \sum_{\substack{n \\ n_1 \equiv (1100) \\ n_2 \equiv n_3 \equiv (0000) \bmod 2}} e^{\frac{2\pi i \tau}{2} \cdot \frac{1}{2} Q_{12}(n)} - \sum_{\substack{n \\ n_4 \equiv (1100) \bmod 2}} e^{\frac{2\pi i \tau}{2} \cdot \frac{1}{2} Q_{12}(n)} \right\}.$$

Daraus folgt

$$\tau_2(n) = \tfrac{1}{8}\left(z'_{12}(n) - z_{12}(n)\right), \tag{31}$$

wo $z'_{12}(n)$ die Anzahl der Lösungen von $\frac{1}{2}Q_{12}(n) = n$ mit $n_1 \equiv (1\,1\,0\,0)$, $n_2 \equiv n_3 \equiv (0\,0\,0\,0) \bmod 2$ ist.

Die benutzten Reihen $\vartheta(\tau; \mathfrak{h}, \mathfrak{A}, P_\nu)$ führten, soweit sie identisch verschwinden, zu Relationen zwischen Darstellungsanzahlen und Darstellungskomponenten. Diese gestatteten es dann, die Koeffizienten $\tau(n)$ und $\tau_2(n)$ durch Anzahlen und Komponenten auszudrücken. Die Komponenten traten in einer speziellen Form auf, die durch die gewählte Kugelfunktion bedingt war. Verallgemeinert man HECKEs Ansatz zur Konstruktion einer Basis von Kugelfunktionen für die Dimension 2 auf die Dimension ν, so erhält man Relationen, die sich von den hier auftretenden bis auf von n unabhängige Faktoren dadurch unterscheiden, daß x_1^n durch ein Monom in $x_1, \ldots, x_{2\nu}$ des Grades ν ersetzt wird.

Hamburg, Mathematisches Seminar der Universität

(Eingegangen am 26. Oktober 1955)

DOLD, A.
Math. Zeitschr. Bd. 65, S. 25—35 (1956)

Erzeugende der Thomschen Algebra $\mathfrak{N}$

Von
ALBRECHT DOLD

Einleitung. In [8] hat R. THOM eine Äquivalenzrelation zwischen kompakten (nicht notwendig zusammenhängenden) differenzierbaren Mannigfaltigkeiten eingeführt, die sich, grob gesprochen, folgendermaßen beschreiben läßt: Zwei differenzierbare Mannigfaltigkeiten sind genau dann äquivalent („cobordantes"), wenn sie zusammen den Rand einer berandeten differenzierbaren Mannigfaltigkeit bilden. (Für eine präzise Definition und für das Folgende vgl. [8], Chap. IV.)

Die Äquivalenzklassen können in natürlicher Weise addiert und multipliziert werden und bilden einen Ring bezüglich dieser Verknüpfungen. Man erhält verschiedene Äquivalenzklassenmengen und verschiedene Ringe Ω oder $\mathfrak{N}$, je nachdem ob man orientierte oder nicht-orientierte Mannigfaltigkeiten betrachtet. Die topologische Dimension der Mannigfaltigkeiten definiert eine Graduierung der Ringe Ω und $\mathfrak{N}$: $\Omega = \sum_{k=0}^{\infty} \Omega^k$, $\mathfrak{N} = \sum_{k=0}^{\infty} \mathfrak{N}^k$.

Die Struktur von $\mathfrak{N}$ ist verhältnismäßig einfach: $\mathfrak{N}$ ist ein Polynomring in abzählbar vielen Variablen x_i über dem Primkörper der Charakteristik 2; x_i wird repräsentiert durch eine geeignet gewählte i-dimensionale Mannigfaltigkeit $P(i)$; i durchläuft alle natürlichen Zahlen, die nicht von der Form $2^l - 1$ sind.

In [8] hat THOM für alle geraden i und für $i = 5$ Mannigfaltigkeiten $P(i)$ angegeben, die die (nicht eindeutig festgelegten) Variablen x_i repräsentieren. In der vorliegenden Arbeit geschieht dies für die verbleibenden Fälle, d.h. für alle ungeraden $i \neq 2^l - 1$, $i > 5$ (s. Satz 3).

Die angegebenen Mannigfaltigkeiten $P(i)$, i ungerade $\neq 2^l - 1$, sind orientierbar. Sie definieren daher Elemente von Ω (und zwar von der Ordnung 2; s. H), die bei der natürlichen Abbildung $\Omega \to \mathfrak{N}$ in x_i übergehen. Dies liefert Aussagen über die nur teilweise bekannte Struktur (s. [8], S. 81) von Ω; z.B. ergibt sich: Ω^k ist nicht trivial für $k \geq 8$, d.h. in jeder Dimension $k \geq 8$ gibt es nicht-berandende orientierbare kompakte differenzierbare Mannigfaltigkeiten (s. H). —

Wir definieren und untersuchen nun zunächst eine Klasse von Mannigfaltigkeiten $\{P(m, n)\}$, unter denen die in der vorstehenden Einleitung genannten Mannigfaltigkeiten $P(i)$ vorkommen.

A. Definition und erste Eigenschaften von $P(m, n)$. S^m, $m \geq 0$, sei die Einheitssphäre des $(m+1)$-dimensionalen euklidischen Raumes mit den Koordinaten $x_0, x_1, \ldots, x_m$. $PC(n)$, $n \geq 0$, sei der komplexe projektive Raum

von n komplexen Dimensionen; wir beschreiben ihn durch homogene Koordinaten $z_0, z_1, \ldots, z_n$.

Der topologische Raum $P(m, n)$[1]) entsteht aus dem topologischen Produkt $S^m \times PC(n)$ durch die Identifikation $(x, z) = (-x, \bar{z})$. Dabei bezeichnet $-x$ den Diametralpunkt von $x \in S^m$ und $\bar{z}$ ist derjenige Punkt aus $PC(n)$, dessen Koordinaten konjugiert komplex zu denen von $z \in PC(n)$ sind.

Bezüglich der Identifikationsabbildung $\Phi: S^m \times PC(n) \to P(m, n)$ ist $S^m \times PC(n)$ zweiblättrige Überlagerung von $P(m, n)$; die nicht-triviale Deckbewegung φ ist durch

$$\varphi(x, z) = (-x, \bar{z}), \qquad x \in S^m,\ z \in PC(n) \tag{1}$$

gegeben.

$S^m \times PC(n)$ kann — in der üblichen Weise — als (reell) analytische Mannigfaltigkeit aufgefaßt werden; φ ist dann eine analytische Abbildung. Vermöge der Abbildung Φ wird daher auch $P(m, n)$ zur (reell) analytischen Mannigfaltigkeit, und Φ selbst ist analytisch.

Die Abbildung $(x, z) \to x$ von $S^m \times PC(n)$ auf S^m geht bei der Identifikation Φ in eine Abbildung $p: P(m, n) \to PR(m)$ ($= m$-dimensionaler reeller projektiver Raum) über. p ist Projektion eines (analytischen) Faserraumes $\{P(m, n), p, PR(m), PC(n), Z_2\}$[2]) (s. [3], § 2); das nicht-triviale Element der Strukturgruppe Z_2 ist die Selbstabbildung $z \to \bar{z}$ von $PC(n)$.

B. Zellenzerlegung und Homologie mod 2 von $P(m, n)$. Wir geben bekannte Zellenzerlegungen von S^m und $PC(n)$ an und gewinnen mit ihrer Hilfe eine Zellenzerlegung von $P(m, n)$.

S^m: Die durch $x_{i-1} = x_{i+2} = \cdots = x_m = 0$, $x_i > 0$ ($x_i < 0$) definierte Punktmenge von S^m ist eine offene i-Zelle C_i^+ (C_i^-). Die Zellen $C_i^\pm$, $i = 0, 1, \ldots, m$, bilden eine Zellenzerlegung von S^m (CW-Komplex im Sinne von J. H. C. WHITEHEAD [10]) und genügen bei geeigneter Orientierung den Berandungsrelationen $\partial C_i^+ = C_{i-1}^+ + C_{i-1}^-$, $\partial C_i^- = -(C_{i-1}^+ + C_{i-1}^-)$, $i = 1, 2, \ldots, m$. Im folgenden seien die Zellen $C_i^\pm$ ein für alle Male mit einer festen Orientierung dieser Art ausgestattet.

Die Diametralpunktvertauschung $x \to -x$, $x \in S^m$, ist eine Selbstabbildung der m-Sphäre, die die Orientierung erhält oder umkehrt, je nachdem ob m ungerade oder gerade ist. Bezüglich der Zellenzerlegung $\{C_i^\pm\}$ ist sie eine Zellenabbildung und führt $C_i^\pm$ in $(-1)^{i+1} C_i^\mp$ über.

***PC(n)*:** Die durch $z_j = 1$, $z_{j+1} = z_{j+2} = \cdots = z_n = 0$ definierte Punktmenge von $PC(n)$ ist eine offene $2j$-Zelle D_j. Die Zellen D_j, $j = 0, 1, \ldots, n$, die wir uns mit einer festen Orientierung versehen denken, bilden eine Zellenzerlegung (CW-Komplex) von $PC(n)$ und genügen den Berandungsrelationen $\partial D_j = 0$.

[1]) $P(1, 2)$ ist die von WU in [12], Nr. 3c betrachtete Mannigfaltigkeit.

[2]) Dies sieht man leicht ein — wir werden es im folgenden nicht benutzen — wenn man beachtet, daß Φ jede offene Menge von der Form $U \times PC(n)$ topologisch abbildet, wo U eine offene Menge aus S^m ist, die keine Diametralpunkte enthält.

Die Selbstabbildungen $z \to \bar{z}$ von $PC(n)$ erhält die Orientierung oder kehrt sie um, je nachdem ob n gerade oder ungerade ist. [Dies erkennt man leicht, wenn man inhomogene Koordinaten $z_1 z_0^{-1}, z_2 z_0^{-1}, \ldots, z_n z_0^{-1}$ benutzt; die Funktionaldeterminante in einem „eigentlichen" Punkt $z_0 \neq 0$ ist $(-1)^n$.] Bezüglich der Zellenzerlegung $\{D_j\}$ ist sie eine Zellenabbildung und führt D_j in $(-1)^j D_j$ über.

$S^m \times PC(n)$: Die Produktzelien $C_i^{\pm} \times D_j$ bilden eine Zellenzerlegung von $S^m \times PC(n)$ und genügen den Berandungsrelationen

$$(2) \qquad \begin{cases} \partial(C_i^{\pm} \times D_j) = \pm\,(C_{i-1}^{+} \times D_j + C_{i-1}^{-} \times D_j) \\ \partial(C_0 \times D_j) = 0, \qquad i = 1, 2, \ldots, m, \qquad j = 0, 1, \ldots, n. \end{cases}$$

Die Selbstabbildung φ [s. (1)] von $S^m \times PC(n)$ ist eine Zellenabbildung und es gilt

$$(3) \qquad \varphi\,(C_i^{\pm} \times D_j) = (-1)^{i+j+1}\,(C_i^{\mp} \times D_j).$$

$P(m, n)$: Die Abbildung Φ: $S^m \times PC(n) \to P(m, n)$ bildet die Zellen $C_i^{\pm} \times D_j$ topologisch ab (vgl. Fußnote 2). Die Zellen $\Phi(C_i^{+} \times D_j)$ — wir bezeichnen sie mit (C_i, D_j) — bilden daher und wegen (3) eine Zellenzerlegung von $P(m, n)$, und Φ ist eine Zellenabbildung. Aus (2) erhält man durch Anwenden von Φ wegen (3) die Berandungsrelationen

$$(4) \qquad \begin{cases} \partial(C_i, D_j) = \left(1 + (-1)^{i+j}\right)(C_{i-1}, D_j) \\ \partial(C_0, D_j) = 0, \qquad i = 1, 2, \ldots, m, \qquad j = 0, 1, \ldots, n. \end{cases}$$

Es bezeichne $H(m, n) = \sum_{\nu} H_\nu(m, n)$ die direkte Summe der Homologiegruppen mod 2, $H^*(m, n) = \sum_{\nu} H^\nu(m, n)$ den Kohomologiering mod 2 von $P(m, n)$. Nach (4) sind alle Zellen (C_i, D_j) Zyklen mod 2. Ihre Homologieklassen $[C_i, D_j]$ bilden daher eine Basis von $H(m, n)$.

$H^*(m, n)$ ist seiner additiven Struktur nach nichts anderes als der Modul der Homomorphismen von $H(m, n)$ in den Primkörper K der Charakteristik 2; $\langle k, h \rangle$ bezeichne den Wert der Kohomologieklasse k auf der Homologieklasse h. Die durch

$$(5) \qquad \langle (c^i, d^j), [C_\mu, D_\nu] \rangle = \delta_\mu^i \cdot \delta_\nu^j, \qquad i, \mu = 0, 1, \ldots, m, \qquad j, \nu = 0, 1, \ldots, n,$$

definierten Kohomologieklassen (c^i, d^j) bilden eine Basis der additiven Struktur von $H^*(m, n)$.

$P(m, n)$ hat also mod 2 dieselben Homologiegruppen wie $PR(m) \times PC(n)$. In D werden wir sehen, daß auch die Kohomologieringe mod 2 dieser beiden Räume übereinstimmen.

C. *$P(m, n)$ ist genau dann orientierbar, wenn $m \neq n\,(2)$ oder $m = 0$ ist.* Denn genau dann ist nach (4) die einzige höchstdimensionale Zelle (C_m, D_n) ein Zykel.

D. Der Kohomologiering mod 2 von $P(m, n)$. Für $m' \leq m$ und $n' \leq n$ identifizieren wir $S^{m'} \times PC(n')$ mit der analytischen Untermannigfaltigkeit

$x_{m'+1}=x_{m'+2}=\cdots=x_m=0$, $z_{n'+1}=z_{n'+2}=\cdots=z_n=0$ von $S^m\times PC(n)$ und $P(m', n')$ mit der analytischen Untermannigfaltigkeit $\Phi(S^{m'}\times PC(n'))$ von $P(m, n)$ (s. A). $P(m', n')$ repräsentiert die Homologieklasse $[C_{m'}, D_{n'}]\in H(m, n)$ (s. B). Wir werden, um den Schnitt der Homologieklassen $[C_i, D_j]$ und $[C_{i'}, D_{j'}]$, $i, i'\leq m$, $j, j'\leq n$, zu bestimmen, die Mannigfaltigkeit $P(i', j')$ in eine andere analytische Untermannigfaltigkeit von $P(m, n)$ deformieren, die sich in allgemeiner Lage relativ zu $P(i, j)$ befindet.

Es seien σ bzw. τ Permutationen der Zahlen 0, 1, 2, ..., m, bzw. 0, 1, 2, ..., n. Durch

$$x_0\to \operatorname{sign}(\sigma)\, x_{\sigma(0)},\quad x_\mu\to x_{\sigma(\mu)},\quad \mu>0, \tag{6}$$

$$z_0\to \operatorname{sign}(\tau)\, z_{\tau(0)},\quad z_\nu\to z_{\tau(\nu)},\quad \nu>0, \tag{7}$$

ist eine analytische Selbstabbildung $(\sigma\times\tau)$ von $S^m\times PC(n)$ gegeben, die mit φ [s. (1)] vertauschbar ist; $(\sigma\times\tau)$ induziert daher eine analytische Selbstabbildung (σ, τ) von $P(m, n)$. Wir verbinden die Matrix der (in den Variablen x_i bzw. z_j linearen) Abbildung (6) bzw. (7) durch eine stetige Schar reeller orthogonaler Matrizen $0(t)$ bzw. $0'(t)$ mit der Einheitsmatrix. $0(t)$ und $0'(t)$ definieren dann eine Schar mit φ vertauschbarer Selbstabbildungen von $S^m\times PC(n)$ und damit eine Schar von Selbstabbildungen von $P(m, n)$, die mit (σ, τ) beginnt und mit der Identität endet. (σ, τ) ist also in die Identität deformierbar und induziert daher den identischen Automorphismus der Homologiegruppe $H(m, n)$.

Um den Schnitt $[C_i, D_j]\circ[C_{i'}, D_{j'}]$ zu bestimmen, betrachten wir die Permutationen

$$\begin{cases}\sigma\colon i'+\mu\to i+1-\mu;\\ \tau\colon j'+\nu\to j+1-\nu,\quad \mu=0,1,\ldots,m,\quad \nu=0,1,\ldots,n.\end{cases} \tag{8}$$

(Falls die in (8) auftretenden Zahlen nicht im Intervall $[0, m]$ bzw. $[0, n]$ liegen, sind sie mod $m+1$ bzw. mod $n+1$ abzuändern.)

Die Mannigfaltigkeit $(\sigma\times\tau)(S^{i'}\times PC(j'))$ wird durch die Gleichungen $x_{i+1-\mu}=0$, $z_{j+1-\nu}=0$, $\mu=1, 2, \ldots, m-i'$, $\nu=1, 2, \ldots, n-j'$, beschrieben. Sie befindet sich relativ zu $S^i\times PC(j)$ in allgemeiner Lage (im Sinne von [1], Ch. VIII, § 2, Nr. 11) und ihr Durchschnitt mit $S^i\times PC(j)$ ist durch $x_{i+i'-m+1}=x_{i+i'-m+2}=\cdots=x_m=0$, $z_{j+j'-n+1}=z_{j+j'-n+2}=\cdots=z_n=0$, gegeben, d.h.

$$(\sigma\times\tau)(S^{i'}\times PC(j'))\cap S^i\times PC(j)=\begin{cases}S^{i+i'-m}\times PC(j+j'-n) & \text{für } i+i'\geq m,\ j+j'\geq n\\ \emptyset & \text{sonst.}\end{cases}$$

Da Φ eine analytische Überlagerung ist, befindet sich auch $(\sigma, \tau)(P(i', j'))$ in allgemeiner Lage relativ zu $P(i, j)$, und es ist

$$(\sigma, \tau)(P(i', j'))\cap P(i, j)=\begin{cases}P(i+i'-m, j+j'-n) & \text{für } i+i'\geq m,\ j+j'\geq n\\ \emptyset & \text{sonst.}\end{cases}$$

Für den Schnitt ihrer Homologieklassen gilt daher (s. [1], Ch. VIII, § 2, Nr. 11)

$$(9)\qquad [C_{i'}, D_{j'}] \circ [C_i, D_j] = \begin{cases} [C_{i+i'-m}, D_{j+j'-n}] & \text{für } i+i' \geq m,\ j+j' \geq n \\ 0 & \text{sonst.} \end{cases}$$

Durch Übergang zum Dualen erhalten wir nun aus (9) die multiplikative Struktur des Kohomologieringes $H^*(m, n)$. Der Wert der zu $[C_{m-i}, D_{n-j}]$ dualen Kohomologieklasse $\Delta[C_{m-i}, D_{n-j}]$ auf $[C_\mu, D_\nu]$, $i, \mu = 0, 1, \ldots, m$, $j, \nu = 0, 1, \ldots, n$, ist gleich der Schnittzahl von $[C_\mu, D_\nu]$ mit $[C_{m-i}, D_{n-j}]$ $(=0$ für $\mu + 2\nu \neq i + 2j)$. Daher folgt aus (5) und (9)

$$(10)\qquad \Delta[C_{m-i}, D_{n-j}] = (c^i, d^j), \qquad i = 0, 1, \ldots, m,\ j = 0, 1, \ldots, n.$$

Satz 1. *Der Kohomologiering* mod 2 *von* $P(m, n)$, $H^*(m, n)$, *wird von den 1- bzw. 2-dimensionalen Klassen* $c = (c^1, d^0)$ *und* $d = (c^0, d^1)$ [s. (5)] *erzeugt. Es bestehen die definierenden Relationen* $c^{m+1} = 0$, $d^{n+1} = 0$.

Bezeichnen c' *und* d' *die entsprechend definierten Erzeugenden von* $H^*(m', n')$, *ist* $m' \leq m$, $n' \leq n$ *und* $f^*: H^*(m, n) \to H^*(m', n')$ *der durch die Injektion* $P(m', n') \to P(m, n)$ (s. D) *induzierte Homomorphismus, so gilt* $f^*(c) = c'$, $f^*(d) = d'$.

Der erste Teil von Satz 1 folgt aus (9) und (10), weil beim Übergang zum Dualen, Δ, die Schnittbildung in $H(m, n)$ in das $\smile$-Produkt in $H^*(m, n)$ übergeht.

Für den zweiten Teil hat man zu beachten, daß $P(m', n')$ nichts anderes ist als die Vereinigung derjenigen Zellen (C_i, D_j) (s. B) von $P(m, n)$, für die $i \leq m'$ und $j \leq n'$ ist; die Behauptung folgt dann aus der Definition von c, c', d, d'.

Bemerkung. Satz 1 läßt sich auch mit Hilfe der spektralen Kohomologiefolge des Faserraums $\{P(m, n), p, PR(m), PC(n), Z_2\}$ (s. A) gewinnen: Man kann die Kohomologie von Basis und Faser als bekannt voraussetzen. Es läßt sich unschwer zeigen, daß die spektrale Kohomologiefolge mod 2 von $P(m, n)$ trivial ist, d.h. daß die Faser mod 2 total nicht nullhomolog ist (s. [2], Chap. III, 7). Dann erhält man die Struktur von $H^*(m, n)$ vermöge [2], Chap. III, Prop. 9.

E. Die „squares" in $P(m, n)$. Die STEENRODschen Sq^i-Kohomologieoperationen[3]) in $H^*(m, n)$ können wegen Satz 1 und wegen der CARTANschen Produktformel $Sq^i(xy) = \sum_{\nu+\mu=i} Sq^\nu x\, Sq^\mu y$ durch die $Sq^\nu c$ und $Sq^\nu d$ ausgedrückt werden. Es gilt

$$(11)\qquad Sq^0 c = c, \quad Sq^1 c = c^2, \quad Sq^i c = 0 \qquad \text{für } i > 1.$$

$$(12)\qquad Sq^0 d = d, \quad Sq^1 d = cd, \quad Sq^2 d = d^2, \quad Sq^j d = 0 \qquad \text{für } j > 2.$$

Während der Kohomologiering mod 2 von $P(m, n)$ nach Satz 1 mit dem von $PR(m) \times PC(n)$ übereinstimmt, fallen die Sq^i nach (12) verschieden aus:

[3]) Mit Sq^i bezeichnen wir diejenige „square"-Operation, mod 2 (s. [4] oder [5]), die die Dimension um i erhöht.

In $PR(m) \times PC(n)$ ist $Sq^1 x = 0$ für jede 2-dimensionale Kohomologieklasse x, in $P(m, n)$ ist $Sq^1 d = cd \neq 0$ für $m, n \geq 1$. Diese Gleichung, $Sq^1 d = cd$, ist die einzige unter den Gln. (11) und (12), die wir beweisen müssen; die anderen sind bekannt (s. [4] oder [5]). Die Kohomologieoperation Sq^1 in $H^*(m, n)$ ist nach [5], Nr. 3 und 4 der zur exakten Folge $0 \to Z_2 \xrightarrow{2} Z_4 \to Z_2 \to 0$ gehörige BOCKSTEIN-Homomorphismus δ^*: $H^*(m, n) \to H^*(m, n)$; Z_r bezeichnet die zyklische Gruppe der Ordnung r. Nach (4) ist $\partial(C_1, D_1) = 2(C_0, D_1)$, $m \geq 1$, $n \geq 1$, und im Rand anderer Zellen als (C_1, D_1) kommt (C_0, D_1) nicht vor. Daraus folgt: $\delta^*(c^0, d^1) = (c^1, d^1)$ [s. (5)], d. h. aber $Sq^1 d = (c^1, d^1) = cd$, letzteres wegen (9) und (10). Diese Gleichung ist auch noch richtig für $m = 0$ oder $n = 0$; beide Seiten sind dann 0 [für $m = 0$, weil dann $H^3(m, n)$ trivial ist].

F. Die STIEFEL-WHITNEYsche Klasse von $P(m, n)$. Ist $\mathfrak{B} = \{B, p, P, S^{r-1}, 0_r\}$ ein Faserraum mit einem Polyeder P als Basis, der $(r-1)$-Sphäre als Faser und der orthogonalen Gruppe 0_r als Strukturgruppe, so bezeichnet $W_j(\mathfrak{B})$ die mod 2 reduzierte j-te STIEFEL-WHITNEYsche Klasse von $\mathfrak{B}$, $j = 0, 1, \ldots, r$ (s. [3], § 38). $W(\mathfrak{B}) = \sum_{j=0}^{r} W_j(\mathfrak{B})$ bezeichnen wir als die STIEFEL-WHITNEYsche Klasse von $\mathfrak{B}$ schlechthin.

Ist insbesondere $P = P(m, n)$ (s. A) und $\mathfrak{B}$ der Faserraum der Tangenteneinheitsvektoren (bezüglich irgendeiner RIEMANNschen Metrik) so schreiben wir $W_j(m, n)$ statt $W_j(\mathfrak{B})$ und $W(m, n)$ statt $W(\mathfrak{B})$.

Satz 2. *Die* STIEFEL-WHITNEY*sche Klasse von $P(m, n)$ ist gegeben durch*

$$W(m, n) = (1 + c)^m (1 + c + d)^{n+1}.$$ [4])

[*Definition von $P(m, n)$ s. A, von c und d s. Satz 1.*]

Beweis durch Induktion nach m und n: Satz 2 ist richtig für $m = n = 0$. Sei nun $m > 0$ oder $n > 0$.

Für $n > 0$ ist $P(m, n-1)$ bezüglich der in D definierten Injektion eine analytische Untermannigfaltigkeit von $P(m, n)$ und repräsentiert die Homologieklasse $[C_m, D_{n-1}] \in H(m, n)$. Die hierzu duale Kohomologieklasse ist $d = (c^0, d^1)$ [s. (10)]. N sei der Faserraum der zu $P(m, n-1)$ normalen Einheitsvektoren von $P(m, n)$. Nach THOM (s. [7] oder [9], Chap. III, Nr. II) gilt

$$\psi^*(W_j(N)) = Sq^j d. \tag{13}$$

Dabei ist ψ^*: $H^*(m, n-1) \to H^*(m, n)$ der zur Injektion $P(m, n-1) \to P(m, n)$ gehörige „Umkehrhomomorphismus"; er genügt der Beziehung (s. [9], Introduct. VI)

$$\psi^*(f^*(y)) = y\, d, \quad y \in H^*(m, n); \tag{14}$$

[4]) Den Satz 2 und damit einen gegenüber dem ursprünglichen erheblich vereinfachten Beweis von Satz 3 verdanke ich Herrn D. PUPPE, der $W(m, n)$ mit der Methode von WU [12] berechnet hat.

$f^*\colon H^*(m, n) \to H^*(m, n-1)$ der durch die Injektion induzierte Homomorphismus.

f^* ist eine Abbildung *auf* (Satz 1); wir wählen $W_i \in H^*(m, n)$ so, daß $f^*(W_i) = W_i(N)$ ist. Aus (13) und (14) folgt dann $W_i d = Sq^i d$. Hieraus und aus (12) $W_0 = 1$, $W_1 \equiv c$, $W_2 \equiv d \bmod d^n$ und somit

$$W(N) = f^*(1 + c + d). \tag{15}$$

Ist T der Faserraum der Tangenteneinheitsvektoren von $P(m, n)$, T' der über $P(m, n-1)$ gelegene Teil von T, so gilt (s. [3], 35.7)

$$W(T') = f^* W(T) = f^* W(m, n). \tag{16}$$

Nach der WHITNEYschen Dualitätsformel (s. [11], § 3) ist

$$W(T') = W(m, n-1)\, W(N). \tag{17}$$

Die Induktionsvoraussetzung über $W(m, n-1)$ besagt wegen Satz 1

$$W(m, n-1) = f^*\big((1 + c)^m (1 + c + d)^n\big). \tag{18}$$

Einsetzen von (15), (16) und (18) in (17) ergibt

$$f^* W(m, n) = f^*\big((1 + c)^m (1 + c + d)^{n+1}\big),$$

d. h.

$$W(m, n) \equiv (1 + c)^m (1 + c + d)^{n+1} \bmod d^n. \tag{19}$$

Diese Formel ist trivialerweise auch richtig für $n = 0$.

Als Zweites betrachten wir die Injektion $P(m-1, n) \to P(m, n)$, falls $m > 0$. $P(m-1, n)$ repräsentiert die Homologieklasse $[C_{m-1}, D_n] \in H(m, n)$ (s. D); hierzu dual ist die Kohomologieklasse $c = (c^1, d^0)$ [s. (10)]. Eine zur vorangehenden genau analoge Betrachtung ergibt nun

$$W(M) = f^*(1 + c) \tag{15'}$$

und

$$W(m, n) \equiv (1 + c)^m (1 + c + d)^{n+1} \bmod c^m. \tag{19'}$$

Dabei bezeichnet $f^*\colon H^*(m, n) \to H^*(m-1, n)$ wieder den durch die Injektion induzierten Homomorphismus und M ist der Faserraum der auf $P(m-1, n)$ normalen Einheitsvektoren von $P(m, n)$.

Nach (19) und (19') ist $W(m, n) \equiv (1+c)^m (1+c+d)^{n+1} \bmod c^m d^n$. Der homogene Bestandteil vom Grade $m + 2n$ auf der rechten Seite dieser Kongruenz ist $(m+1)(n+1) c^m d^n$. Andererseits ist (s. [3], 39.7) $W_{m+2n}(m,n) = N(m, n) c^m d^n$, wenn $N(m, n)$ die EULERsche Charakteristik von $P(m, n)$ bezeichnet. Der Beweis des Satzes 2 ist also vollendet, wenn wir noch gezeigt haben, daß $N(m, n) \equiv (m+1)(n+1) \bmod 2$ ist. Aus der Struktur von $H^*(m, n)$ (s. Satz 1) ergibt sich aber durch einfache Rechnung

$$N(m, n) = \tfrac{1}{2}\big(1 + (-1)^m\big)(n + 1) \equiv (m + 1)(n + 1) \bmod 2.$$

G. Die Erzeugenden von $\mathfrak{N}$.

Satz 3. *Für jede natürliche Zahl i, die nicht von der Form 2^l-1 ist, seien die ganzen Zahlen r und s durch $i+1=2^r(2s+1)$ erklärt. Es sei*

$$P(i)=\begin{cases} P(i,0)=PR(i) & \text{für gerades } i \\ P(2^r-1, s2^r) & \text{für ungerades } i \text{ (s. A).}\end{cases}$$

$P(i)$ ist eine i-dimensionale kompakte differenzierbare Mannigfaltigkeit; x_i bezeichne ihre Klasse in der Algebra $\mathfrak{N}$ (s. [8], Chap. IV).

$\mathfrak{N}$ kann aufgefaßt werden als Polynomalgebra $K[x_2, x_4, x_5, \ldots]$ in den Variablen x_i, $i=2,4,5,6,8,\ldots$ über dem Primkörper K der Charakteristik 2.

NB. Satz 3 geht nur insofern über das Ergebnis von THOM (s. [8], S. 79ff.) hinaus, als auch für die ungeraden Dimensionen $i>5$ explizit repräsentierende Mannigfaltigkeiten für die Variablen x_i angegeben werden.

Beweis. Es seien $K[\iota]$ der Ring der Polynome über K in den Variablen $t_1, t_2, \ldots, t_\iota$, $S[\iota]\subset K[\iota]$ die Unteralgebra der symmetrischen Polynome, $S_j(\iota)$ die j-te elementarsymmetrische Funktion der t_ν, $j=0,1,\ldots,\iota$, und $S^h(\iota)=\sum_{\nu=1}^{\iota} t_\nu^h$, $h=0,1,\ldots$.

Für jedes $\iota\geq i$ definieren wir einen gradtreuen Homomorphismus $\Psi_{i,\iota}$: $S[\iota]\to H^*(i)$ durch

$$\Psi_{i,\iota}(S_j(\iota))=\begin{cases} W_j(i) & \text{für } j\leq i \\ 0 & \text{für } j>i.\end{cases} \tag{20}$$

Dabei bezeichnet $H^*(i)$ den Kohomologiering mod 2 und $W_j(i)$ die j-te STIEFEL-WHITNEYsche Klasse von $P(i)$.

$\Psi_{i,\iota}$ führt $\prod_{\nu=1}^{\iota}(1+t_\nu)=\sum_{\nu=0}^{\iota} S_\nu(\iota)$ in $W(i)=\sum_{\nu=0}^{i} W_\nu(i)$ über und ist unter den gradtreuen Homomorphismen $S[\iota]\to H^*(i)$ dadurch charakterisiert.

Für den Beweis des Satzes 3 genügt es nach [8], Chap. IV, Nr. 7 zu zeigen, daß $\Psi_{i,i}(S^i(i))\neq 0$ ist; wiederum nach [8] können wir uns dabei auf ungerades i beschränken. Wir beweisen nun zunächst die Gleichung

$$\Psi_{i,\iota}(S^h(\iota))=\Psi_{i,i}(S^h(i)),\quad \iota\geq i \tag{21}$$

und zeigen dann $\Psi_{i,i+2}(S^i(i+2))\neq 0$.

Um (21) zu erhalten, definieren wir einen gradtreuen Homomorphismus $p_{i,\iota}$: $K[\iota]\to K[i]$ durch $p_{i,\iota}(t_\nu)=\begin{cases} t_\nu & \text{für } \nu\leq i \\ 0 & \text{für } \nu>i\end{cases}$. Man bestätigt leicht, daß

$$p_{i,\iota}(S_j(\iota))=\begin{cases} S_j(i) & \text{für } j\leq i \\ 0 & \text{für } j>i\end{cases} \tag{22}$$

und

$$p_{i,\iota}(S^h(\iota))=S^h(i) \tag{23}$$

ist. Aus (22) und (20) folgt $\Psi_{i,i}\circ p_{i,\iota}(S_j(\iota))=\Psi_{i,\iota}(S_j(\iota))$, $j=0,1,\ldots,\iota$, also $\Psi_{i,i}\circ(p_{i,\iota}|S[\iota])=\Psi_{i,\iota}$. Wendet man auf (23) $\Psi_{i,i}$ an, so folgt hieraus (21).

Für die Berechnung von $\Psi_{i,i+2}(S^i(i+2))$ benötigen wir die Darstellung von $S^h(2)$ als Polynom in $S_1(2)$ und $S_2(2)$. Die bekannte Identität

$$\sum_{j=0}^{i} S_j(\iota)\, S^{h-j}(\iota) = 0, \quad h \geqq \iota \tag{24}$$

[Beweis: Das Polynom $\prod_{\nu=1}^{i}(1+t t_\nu) = \sum_{j=0}^{i} S_j(\iota) t^j$ über $K[\iota]$ hat die Nullstellen t_ν^{-1}, $\nu = 1, 2, \ldots, \iota$. Man erhält also, wenn man t_ν^{-1} in dieses Polynom einsetzt und mit t_ν^h multipliziert: $0 = \sum_{j=0}^{i} S_j(\iota)\, t_\nu^{h-j}$ und hieraus durch Summation über ν die Gl. (24).]

ergibt für $\iota = 2$

$$S^h(2) = S_1(2)\, S^{h-1}(2) + S_2(2)\, S^{h-2}(2), \quad h \geqq 2. \tag{25}$$

Hieraus folgt durch Induktion nach h

$$S^h(2) = \sum_{p+2q=h} b(p+q-1, q)\, (S_1(2))^p\, (S_2(2))^q, \quad h = 1, 2, \ldots. \tag{26}$$

Dabei ist $b(p+q-1, q)$ der Binomialkoeffizient $\binom{p+q-1}{q} \bmod 2$ ($=1$ für $q=0$, $=0$ für $q<0$ oder $p<q+0$), und es wird über alle nicht negativen ganzen p, q mit $p+2q=h$ summiert.

(26) ist richtig für $h=1, 2$, denn $S^1(2) = S_1(2)$ und $S^2(2) = (S_1(2))^2$. Für $h>2$ ergeben (25) und die Induktionsvoraussetzung

$$\begin{aligned} S^h(2) &= \sum_{p+2q=h-1} b(p+q-1, q)\, (S_1(2))^{p+1}\, (S_2(2))^q + \\ &+ \sum_{p+2q=h-2} b(p+q-1, q)\, (S_1(2))^p\, (S_2(2))^{q+1} \\ &= \sum_{p+2q=h} \big(b(p+q-2, q) + b(p+q-2, q-1)\big)\, ((S_1(2))^p\, (S_2(2))^q \\ &= \sum_{p+2q=h} b(p+q-1, q)\, ((S_1(2))^p\, (S_2(2))^q, \end{aligned}$$

letzteres wegen der bekannten Additionsformel für Binomialkoeffizienten.

Durch die Festsetzungen

$$\left.\begin{aligned} \Psi(t_{2\nu-1} + t_{2\nu}) &= c \\ \Psi(t_{2\nu-1} t_{2\nu}) &= d \\ \Psi(t_{2n+2+\mu}) &= c \end{aligned}\right\} \quad \begin{aligned} &\nu = 1, 2, \ldots, n+1, \quad \mu = 1, 2, \ldots, m, \\ &n = s2^r, \quad m = 2^r - 1 \quad \text{(s. Satz 3)} \end{aligned}$$

definieren wir einen gradtreuen Homomorphismus der von den $(t_{2\nu-1}+t_{2\nu})$, $(t_{2\nu-1}t_{2\nu})$ und $t_{2n+2+\mu}$ erzeugten Unteralgebra von $K[i+2]$ in $H^*(i) = H^*(m, n)$. Ψ führt das Polynom

$$\begin{aligned} \prod_{\varrho=1}^{i+2}(1+t_\varrho) &= \prod_{\nu=1}^{n+1}(1+t_{2\nu-1})\,(1+t_{2\nu}) \prod_{\mu=1}^{m}(1+t_{2n+2+\mu}) \\ &= \prod_{\nu=1}^{n+1}\big(1+(t_{2\nu-1}+t_{2\nu})+t_{2\nu-1}t_{2\nu}\big) \cdot \prod_{\mu=1}^{m}(1+t_{2n+2+\mu}) \end{aligned}$$

in $(1+c+d)^{n+1}(1+c)^m = W(m,n) = W(i)$ über. Die Einschränkung von Ψ auf $S[i+2]$ stimmt also mit $\Psi_{i,i+2}$ überein. Daher ist

$$\begin{aligned}\Psi_{i,i}(S^i(i)) &= \Psi_{i,i+2}(S^i(i+2)) = \Psi(S^i(i+2)) = \Psi\Big(\sum_{\varrho=1}^{i+2} t_\varrho^i\Big)\\ &= (n+1)\,\Psi(t_1^i + t_2^i) + m\,\Psi(t_{i+2}^i) = (n+1)\,\Psi(S^i(2)) + m\,\Psi(t_{i+2}^i)\\ &= (n+1)\,\Psi\Big(\sum_{p+2q=i} b(p+q-1,q)\,(S_1(2))^p\,(S_2(2))^q\Big) + m\,\Psi(t_{i+2}^i)\\ &= (n+1)\sum_{p+2q=i} b(p+q-1,q)\,c^p d^q + m\,c^i.\end{aligned}$$

Wegen $i > m$ ist $c^i = 0$ (s. Satz 1). Ferner ist $c^p d^q$ für $p+2q=i$ nur dann von 0 verschieden, wenn $p=m$ und $q=n$ ist (Satz 1). Beachtet man noch, daß $n+1 \equiv 1(2)$ ist, so ergibt sich $\Psi_{i,i}(S^i(i)) = b(m+n-1,n)\,c^m d^n$. Nun ist $m+n-1 = s2^r + (2^r-2)$ und $n = s2^r$. Daraus geht hervor, daß die dyadische Entwicklung von n in der von $m+n-1$ enthalten ist, d.h. (s. [6], Lemma 2.2), daß $b(m+n-1,n) = 1$ und $\Psi_{i,i}(S^i(i)) = c^m d^n \neq 0$ ist, w.z.b.w.

Bemerkung. Die in Satz 3 getroffene Auswahl der Erzeugenden von $\mathfrak{N}$ unter den Mannigfaltigkeiten $P(m,n)$ ist nicht eindeutig festgelegt. Aus dem Beweis geht hervor, daß man z.B. in der Dimension 6 auch $P(2,2)$ statt $P(6,0)$ und in der Dimension 9 $P(5,2)$ statt $P(1,4)$ hätte nehmen können.

H. Anwendung auf die Algebra Ω. Für ungerades $i \neq 2^j-1$ ist die Mannigfaltigkeit $P(i)$ (s. Satz 3) orientierbar (s. C). Versehen wir sie mit einer festen Orientierung, so definiert sie ein Element $v_i \in \Omega^i$ (s. Einleitung oder [8], Chap. IV, 1), das beim natürlichen Homomorphismus $\Omega \to \mathfrak{N}$ in x_i übergeht. v_i ist also sicher nicht Null. Es gilt aber: *Die Ordnung von $v_i \in \Omega^i$ ist* 2.

Es gibt nämlich eine differenzierbare, orientierungsumkehrende Involution von $P(i)$. Um dies einzusehen, betrachten wir das Produkt $S^m \times PC(n)$ mit $m = 2^r-1$, $n = s2^r$ (s. Satz 3). Ist $x = (x_0, x_1, \ldots, x_m)$ ein Punkt aus S^m, so bezeichne $\tilde{x}$ den Punkt $(x_0, x_1, \ldots, x_{m-1}, -x_m)$ aus S^m. Durch $\tilde{I}(x,z) = (\tilde{x}, z)$, $x \in S^m$, $z \in PC(n)$, definieren wir eine differenzierbare Involution von $S^m \times PC(n)$ vom Grade -1. $\tilde{I}$ ist mit der Deckbewegung φ [s. (1)] vertauschbar und induziert daher eine differenzierbare Involution I von $P(m,n)$; I kann durch die Gleichung $I \circ \Phi = \Phi \circ \tilde{I}$ definiert werden. Φ hat als zweiblättrige Überlagerung den Grad ± 2, also I wegen $I \circ \Phi = \Phi \circ \tilde{I}$ den Grad -1.

Hieraus folgt nun: *Beim natürlichen Homomorphismus $\Omega \to \mathfrak{N}$ ist die von den x_i (s. Satz 3) mit ungeradem $i \neq 2^j-1$ erzeugte Unteralgebra von $\mathfrak{N}$ isomorphes Bild der von den v_i erzeugten Unteralgebra von Ω.* Insbesondere ergibt sich, daß die Gruppen Ω^k nicht zu „klein" sein können; z.B. gilt: *Für $k \geq 8$ ist Ω^k nicht trivial*, d.h. in jeder Dimension $k \geq 8$ gibt es nicht-berandende kompakte orientierbare differenzierbare Mannigfaltigkeiten.

Zum Beweis genügt es (nach dem eben Gesagten und nach [8], Chap. IV, Nr. 8), ein Produkt von Räumen $P(i)$, i ungerade $\neq 2^j-1$, und komplexen projektiven Räumen $PC(2j)$ gerader komplexer Dimension anzugeben, das

die Gesamtdimension k hat; d.h. wir haben eine Partition von k in ungerade natürliche Zahlen $\neq 2^l-1$ und durch 4 teilbare natürliche Zahlen anzugeben. Für $k \equiv 3\,(4)$, $k \geq 11$, ist $k = 11 + (k-11)$, für $k \equiv 2\,(4)$, $k \geq 10$, ist $k = 5+5+(k-10)$ und in allen anderen Fällen ist $k = k$ eine solche Partition.

Literatur

[1] Lefschetz, S.: Topology. Amer. Math. Soc. Colloq. Publ. **12** (1930). — [2] Serre, J. P.: Homologie singulière des espaces fibrés; applications. Ann. of Math. **54**, 425—505 (1951). — [3] Steenrod, N.: The topology of fibre bundles. Princeton Math. Ser. 14. — [4] Steenrod, N.: Products of cocycles and extensions of mappings. Ann. of Math. **48**, 290—320 (1947). — [5] Steenrod, N.: Homology groups of symmetric groups and reduced power operations; cyclic reduced powers of cohomology classes. Proc. Nat. Acad. Sci. U.S.A. **39**, 213—223 (1953). — [6] Steenrod, N., and J. H. C. Whitehead: Vector fields on the n-sphere. Proc. Nat. Acad. Sci. U.S.A. **37**, 58—63 (1951). — [7] Thom, R.: Variétés plongées et i-carrés. C. R. Acad. Sci. Paris **230**, 507—508 (1950). — [8] Thom, R.: Quelques propriétés globales des variétés différentiables. Comment. Math. Helv. **28**, 17—87 (1954). — [9] Thom, R.: Espaces fibrés en sphères et carrés de Steenrod. Ann. Ecole norm. sup. (3) **69**, 109—182 (1952). — [10] Whitehead, J. H. C.: Combinatorial Homotopy. I. Bull. Amer. Math. Soc. **55**, 213—245 (1949). — [11] Wu, Wen-Tsün: On the product of sphere-bundles and the duality theorem mod 2. Ann. of Math. **49**, 641—653 (1948). — [12] Wu, Wen-Tsün: Classes caractéristiques et i-carrés. C. R. Acad. Sci. Paris **230**, 508—511 (1950).

Heidelberg, Mathematisches Institut der Universität

(Eingegangen am 22. November 1955)

NEUMANN, H.
Math. Zeitschr. Bd. 65, S. 36—69 (1956)

On varieties of groups and their associated near-rings

By
HANNA NEUMANN

Introduction

The present paper arose out of the attempt to clarify some questions on varieties of groups and laws in groups. The central theme is the role played by the number of variables involved in laws. For reasons that will appear later, certain near-rings connected with free groups proved a helpful tool. Once introduced, these near-rings raise their own problems. The work done on near-rings by previous authors (cf. e.g. [2], [3]) has very little bearing on those arising here: we have no chain conditions available; but to a certain extent the similarity of our near-rings to ordinary matrix rings makes up for this.

The first part (§§ 1—3) deals with the interrelation between certain varieties of groups. A variety of groups is determined by the set of all laws holding in all groups of the variety or, equivalently, by a fully invariant subgroup V of the free group $F = \{e_1, e_2, \ldots\}$ of countably infinite rank. The intersection V_n of V with the subgroup F_n of F which is generated by $e_1, e_2, \ldots, e_n$, determines in the form F_n/V_n the free group of rank n of the variety. A number of questions arise naturally in this context; for example we may ask for the different varieties of which F_n/V_n is the free group of rank n. Or again, bearing in mind that V_n represents the laws in n variables of the variety determined by V, we enquire into the relation between $V_m = F_m \cap V$ and the set of all laws in m variables that hold in F_n/V_n. The results may be summarized as follows: Every variety, that is every fully invariant subgroup V of F, determines two chains of fully invariant subgroups, one descending and one ascending, both "converging" to V. These two chains are in a sense dual to each other; while the nth term of the ascending chain corresponds to the set of all laws in n variables of the variety, the nth term of the descending chain represents the set of all laws holding in every group of the variety that can be generated by n elements. The group F_n/V_n is the free group of rank n of precisely all those varieties which are determined by fully invariant subgroups of F "between" the nth term of the ascending chain and the nth term of the decending chain. Several of the many unsolved problems on laws in groups are equivalent to the question whether one or the other of these chains becomes stationary after a finite number of steps, or is stationary in certain places.

The situation described here is probably widely known. But while certain aspects of it, and specific examples, can be found in the literature (cf. e.g. [8], [9], [11], [14]), no connected account seems to exist. This is the justification of the first part: what little we use of it in the second part, mainly the significance of the ascending chain, could have been dealt with rather more briefly.

Part II is devoted to the near-rings mentioned before. They are introduced in § 4. We consider a reduced free group G_n of finite or infinite rank n, that is a group which is the free group of rank n of some variety, generated freely by the elements $e_1, e_2, \ldots, e_n$. These generators have the property (characteristic for reduced free groups[1])), that every mapping of them into the group can be continued to an endomorphism of the group. Consequently addition of endomorphisms may be so defined that, together with the usual definition of multiplication, the endomorphisms of G_n form a near-ring Γ_n. The definition of addition differs from that usual for mappings of a group into itself ([2], [15]), with the result that in our near-rings the right distributive law holds, that is a right-hand factor distributes over a sum as left-hand factor, while in the near-ring of all mappings of a group into itself one customarily has the left distributive law[2]).

The representation of an endomorphism by means of its n "components", namely the maps of the free generators, leads to an interpretation of these near-rings as a generalization of matrix rings. In fact if G_n is the free group of an abelian variety, this interpretation coincides with the representation of the ring of endomorphisms as ring of matrices over the integers (or integers modulo m). A matrix ring possesses a module basis consisting of those matrices having 1 in the ith row and j^{th} column and zeros elsewhere. Our near-rings contain corresponding elements; they also form an additive basis, if n is finite. In addition they have the property that they distribute on both sides. This "distributive basis" plays a decisive part. However, the basis property gets lost if n is infinite. For this reason we confine ourselves subsequently to finite and countably infinite rank, and in the latter case restrict ourselves to what may be loosely described as the near-ring of all row- and column-finite matrices (§ 5). This near-ring, denoted by Γ_ω, can be represented as the union of all near-rings Γ_n for finite n and therefore still possesses the basis property. Due to the existence of a "distributive basis", the ideals in the near-rings Γ_n and Γ_ω, defined as normal subgroups of the additive group which are closed under right and left multiplication by elements of the near-ring, are precisely the kernels of the homomorphisms of these near-rings (§ 6). On the other hand, the specific nature of our near-rings permits a different characterization of an ideal, namely as the subset of all those matrices whose components vary over a fixed fully invariant subgroup of the associated reduced free group. This remains true for the restricted near-ring Γ_ω in place of the near-ring of all endomorphisms of the reduced free group G of countably infinite rank; and this is the reason that these near-rings Γ_ω are sufficient for our purposes. This characterization then shows that all near-rings Γ_n and Γ_ω are homomorphic images of the corresponding near-rings Φ_n and Φ_ω associated with the absolutely free groups F_n and F respectively. Next we discuss the formal properties of the

[1]) Cf. [6]; P. HALL uses the term "relatively free group".

[2]) In ZASSENHAUS however this becomes the right distributive law due to his definition of the product $\alpha\beta$ as the mapping β followed by the mapping α.

lattice of ideals and introduce ideal multiplication (§ 7). The similarity to matrix rings is essential here. Ideal multiplication in the near-ring corresponds in the group to the familiar process of obtaining new fully invariant subgroups by substituting the elements of one for the generators in the words of another fully invariant subgroup. It is this process of substitution in the free group which benefits most by being translated into multiplication in the near-ring. The limitation of the number of variables for which one may substitute in the case of finite rank n then expresses itself in the fact that ideal multiplication in Γ_n is in general not associative. By contrast, ideal multiplication proves associative in the near-rings Γ_ω. In the last two paragraphs we specialize to the near-rings Φ_n and Φ_ω belonging to the absolutely free groups, mainly because we make use of the SCHREIER property that every subgroup of a free group is itself free. The ideals in Φ_ω (§ 8) form a semigroup with right cancellation; in fact more than that, right multiplication by a fixed ideal is an iso-endomorphism of the lattice of ideals. The fully invariant subgroups of F which are fully invariant in a larger fully invariant subgroup of F are just those corresponding to ideal products in Φ_ω. This is no longer true in the case of finite rank n (§ 9): here they correspond merely to the intersections of ideal products in Φ_ω with the near-ring Φ_n. A closer analysis of the situation shows this to be the immediate reason for the non-associativity of ideal multiplication in Φ_n.

In conclusion it should be mentioned that our representation of the endomorphisms of free groups as a near-ring of matrices can be carried through similarly for the free algebras of more general varieties, for example the varieties of loops, or of groupoids. In fact, precisely this has been done by A. ROBINSON—in different language—for the free groupoid on one generator (cf. [13]; compare also the logarithmetics of I. M. H. ETHERINGTON: cf. [4] and the literature there quoted). ROBINSON states (loc. cit. p. 117) that there is no "natural" definition of multiplication in terms of addition in the case of more than one generator. However, our definition of multiplication of n-rowed "matrices" will do just that also in the case of a free groupoid on n generators.

Part I

1. Preliminaries and notation

We begin with a summary of definitions and facts that will be used throughout, introducing at the same time the notation employed.

All groups will be written additively, the neutral element zero will be denoted by 0 in every group. Capital Roman letters signify groups, small Roman letters denote elements of groups. Small Greek letters stand for endomorphisms of groups. We write $2x$ for $x+x$, generally $nx = x+x+\cdots$, where the sum has n terms; $[x, y]$ denotes the commutator of x and y, $[x, y] = -x-y+x+y$; $[x, y, z]$ is the left-normed commutator, $[x, y, z] = [[x, y], z]$.

The free group on $\aleph_0$ generators $e_1, e_2, \ldots$ is denoted by F, the subgroup generated by $e_1, e_2, \ldots, e_n$, for every finite n, is denoted by F_n. Then $F = \bigcup_n F_n$.

9272-65

F_n may be obtained from F by means of the retractive homomorphism ε_n defined by $e_i\varepsilon_n = e_i$ for $i = 1, \ldots, n$ and $e_i\varepsilon_n = 0$ for $i > n$. Then $F_n = F\varepsilon_n \cong F/F_n^*$, where F_n^*, the kernel of the homomorphism, is the normal closure in F of the free factor generated by $e_{n+1}, e_{n+2}, \ldots$.

1.01. *An element w of F belongs to F_n if, and only if, $w\varepsilon_n = w$. If V is a fully invariant subgroup of F, its intersection V_n with F_n is given by $V_n = V \cap F_n = V\varepsilon_n$. V_n is a fully invariant subgroup of F_n, and $V = \bigcup_n V_n$.*

A *variety of groups* is the class of all groups satisfying a given set of laws (or identical relations). Here a law is a word $v(x_1, x_2, \ldots)$ in the variables x_i; the law holds in the group H if for every substitution of elements $h_1, h_2, \ldots$ of H for the variables $x_1, x_2, \ldots$ one has $v(h_1, h_2, \ldots) = 0$ qua element of H. If $v(x_1, x_2, \ldots)$ varies over the set of all laws satisfied by the groups of the variety, the corresponding words $v(e_1, e_2, \ldots)$ form a fully invariant subgroup V of F. We say: V represents the laws of the variety, the variety corresponds to, or is determined by, the fully invariant subgroup V of F. Every fully invariant subgroup V of F determines a variety:

1.02. *A group H belongs to the variety determined by V if, and only if, every countable subgroup of H is a homomorphic image of F/V. If in particular H can be generated by n elements, then H is a homomorphic image of F_n/V_n.*

Since a variety is determined by all its countable groups, we shall henceforth consider countable groups only.

1.03. *The free groups of rank $\aleph_0$ and n respectively of the variety determined by V are given by $G \cong F/V$ and $G_n \cong F_n/V_n$.*

The letters G and G_n will be reserved for *reduced free groups*, that is groups which are the free groups of some variety.

1.04. *A group is reduced free if, and only if, it possesses a set of generators whose every mapping into the group can be continued to an endomorphism of the group.*

Such a set of generators is called a set of free generators (or canonical generators; cf. [6]). In particular, the maps of the generators e_i of F (or F_n) under the natural homomorphism of F (or F_n) onto G (or G_n) according to 1.03 form a set of free generators of G (or G_n).

We use the following terminology:

1.05. Definition. *For any group H, the variety defined by H is the variety corresponding to the set of all laws that hold in H.*

As H is countable, it may be represented as a factor group of F, $H \cong F/R$; if H can be generated by n elements, then also $H \cong F_n/R_n$. Here R, R_n are not, in general, uniquely determined by H; but in both cases the greatest fully invariant subgroup of F or F_n contained in R or R_n respectively depends only on H. One has (cf. [11]):

1.06. *The variety defined by H is the variety corresponding to the greatest fully invariant subgroup V of F contained in R. The set of all laws of H involving*

at most n variables is given by $V_n = V \cap F_n$. *If* $H \cong F_n/R_n$ *for some normal subgroup* R_n *of* F_n, *then* V_n *is the greatest fully invariant subgroup of* F_n *contained in* R_n.

Different groups may, of course, define the same variety. The variety corresponding to V is, trivially, the variety defined by F/V.

For any set of elements of F (or F_n), the least fully invariant subgroup of F (or F_n) containing the set is called the *fully invariant closure in* F (or *in* F_n) of the given set of elements. It is obtained by forming the closure of the set with respect to the operations of addition, and of substitution of arbitrary elements of F (or F_n) for the generators. The set of elements is referred to as a *basis for* V *in* F (or *for* V_n *in* F_n).

2. The ascending chain determined by a variety

Our object is to examine the relation between the fully invariant subgroups of the groups F_n of finite rank and those of the group F of infinite rank.

2.01. Definition. *If* V_m *is a fully invariant sugbroup of* F_m, *the extension of* V_m *to* F_n $(n > m)$, *or to* F, *is the fully invariant closure of* V_m *in* F_n, *or in* F, *respectively.*

If V, *or* V_n, *is a fully invariant subgroup of* F, *or* F_n, *the restriction of* V, *or* V_n, *to* F_m *is the intersection* $V \cap F_m$, *or* $V_n \cap F_m$, *respectively.*

These two processes of extension and restriction are not in general inverse to each other. While an extension followed by restriction of the result to the previous "level" leads back to the original fully invariant subgroup, this is not always true if the processes are carried out in reverse order.

We first note some obvious facts:

2.02. *If* $l < m < n$, *extending* V_l *first to* F_m, *then extending the result to* F_n, *amounts to the same as extending* V_l *to* F_n. *The same is true with* F *in place of* F_n.

2.03. *The extension of* V_m *to* F *is the union of the extensions of* V_m *to* F_n *for* $n = m+1, m+2, \ldots$.

Corresponding to 2.02 we have trivially:

2.04. *If* $l < m < n$, *restricting* V_n *first to* F_m, *then restricting the result to* F_l, *amounts to the same as restricting* V_n *to* F_l. *The same is true with* F *in place of* F_n.

2.05. *If* V_n *is a fully invariant subgroup of* F_n, $V^{(n)}$ *its extension to* F, *then* $V^{(n)} \cap F_n = V_n$.

Proof. Clearly $V_n \subseteq V^{(n)} \cap F_n$. To prove the reverse inclusion, let v be an arbitrary element of $V^{(n)}$ that belongs to F_n. As $V^{(n)}$ is the fully invariant closure of V_n in F, the element v is of the form

$$v = \sum_i v_i(x_{i,1}, \ldots, x_{i,n}),$$

where $v_i(e_1, \ldots, e_n) \in V_n$, and the $x_{i,j}$ are elements of F, that is representable as words $x_{i,j} = x_{i,j}(e_1, e_2, \ldots, e_{k(i,j)})$. But $v \in F_n$, hence $v = v\varepsilon_n$; therefore

$$v = v\varepsilon_n = \sum_i v_i(x_{i,1}\varepsilon_n, \ldots, x_{i,n}\varepsilon_n).$$

But $x_{i,j}\varepsilon_n$ lies in F_n for all i and j, and so each term of this sum belongs to V_n, that is $v \in V_n$.

2.05 combined with 2.02 and 2.04 gives at once:

2.06. *If $l < m < n$, extending V_l to F_n, then restricting the result to F_m, amounts to the same as extending V_l to F_m. The same is true with F in place of F_n.*

We now examine the result of restricting first, and then extending. In this context the following notation will be used throughout the paper:

Notation. If V is a fully invariant subgroup of F, then V_n denotes its restriction to F_n, $V_n = V \cap F_n$; and $V^{(n)}$ is the extension of V_n to F.

Since $V_n \leqq V \cap F_{n+1} = V_{n+1}$, it follows that $V^{(n)} \leqq V^{(n+1)}$ for all n. Moreover, as V is the union of its restrictions V_n for all n, V is a fortiori the union of the groups $V^{(n)}$ for all n. Therefore:

2.07. *The fully invariant subgroups $V^{(n)}$ $(n = 1, 2, \ldots)$ of F determined by V form an ascending chain, and V is the union of its terms;*

$$V^{(1)} \leqq V^{(2)} \leqq \cdots \leqq V^{(n)} \leqq \cdots \leqq V; \qquad V = \bigcup_n V^{(n)}.$$

The fact that $V^{(1)}$ is the trivial group whenever V is contained in the derived group of F shows that the inclusion may be proper for some terms. We shall give a less trivial example at the end of the paragraph. The significance of the chain is explained by:

2.08. Theorem. *The ascending chain determined by V is finite, i.e. it becomes stationary after a finite number of steps if, and only if, V possesses a basis whose members involve only a finite number of generators. If k is the least integer such that V is the closure of a set of words each of which involves at most k generators, then*

$$V^{(k-1)} < V^{(k)} = V^{(k+1)} = \cdots = V.$$

Corollary. *If V is finitely based, the ascending chain is finite.*

Proof. If for some n, $V^{(n)} = V$, then V is the closure of V_n in F, that is, it possesses a basis involving at most $e_1, \ldots, e_n$; thus $n \geqq k$. Conversely, if the words $v_i(e_1, \ldots, e_{n(i)})$ with $n(i) \leq k$ form a basis of V, then V is the closure of V_k in F, as the basis elements all belong to F_k, hence to V_k.

The premise of the corollary touches upon a central unsolved problem, namely the question whether every fully invariant subgroup of F is finitely based, or equivalently, whether the ascending chain condition for fully invariant subgroups holds in F. We cannot even state the converse of the corollary, for it is not known either, whether every fully invariant subgroup of a free group of finite rank is finitely based. If V represents the set of all laws of a finite group, then all factor groups F_n/V_n are finite (cf. [11]), hence by SCHREIER's Theorem V_n as free subgroup of F_n is of finite rank, hence it is certainly finitely based. Therefore we have in that case:

2.09. *If V represents the set of laws of a finite group, the ascending chain is finite if, and only if, V is finitely based.*

Again it is not known, whether there exists a finite group whose laws are not finitely based. R. C. LYNDON [9] has shown that the set of all laws of every finite nilpotent group is finitely based.

Finally we determine the restrictions of the terms of the chain:

2.10. Theorem. *If k is defined as in* 2.08, *then*

$$V^{(n)} \cap F_m = V \cap F_m = V_m, \quad \text{if } n \geq m,$$
$$V^{(n)} \cap F_m = V_m, \quad \text{if } k \leq n < m,$$
$$V^{(n)} \cap F_m < V_m, \quad \text{if } n < k \leq m.$$

Proof. By 2.05, $V^{(m)} \cap F_m = V_m$; also by the definition of V_m, $V \cap F_m = V_m$. If $n \geq m$, then $V^{(m)} \leqq V^{(n)} \leqq V$, hense also $V^{(n)} \cap F_m = V_m$.

Next, $k \leq n < m$ implies $V^{(k)} = V^{(n)} = V^{(m)}$, by definition of k; hence $V^{(n)} \cap F_m = V^{(m)} \cap F_m = V_m$.

Finally, $n < k \leq m$ implies $V^{(n)} \leqq V^{(k-1)} < V^{(k)} = V^{(m)}$; hence $V^{(n)} \cap F_m \leqq V^{(m)} \cap F_m = V_m$. By 2.06, $V^{(n)} \cap F_m$ is the extension of V_n to F_m, therefore, by 2.02, the extension of $V^{(n)} \cap F_m$ to F is $V^{(n)}$. If now we had $V^{(n)} \cap F_m = V_m$, then $V^{(n)} = V^{(m)}$ would follow, since the extension of V_m to F is $V^{(m)}$. Hence $V^{(n)} \cap F_m < V_m$.

We conclude the paragraph with an illustration:

2.11. Example. We consider the second derived group of F, which we denote by A for the present[3]). Thus A is seen to be the closure in F of the word $[[e_1, e_2], [e_3, e_4]]$, and therefore $k \leq 4$. Although $k = 4$ is plausible, we know at present only:

$$A^{(1)} = \{0\} < A^{(2)} \leqq A^{(3)} \leqq A^{(4)} = A.$$

The restriction of every term except the first to the free group F_2 is, by 2.10, the same group, namely the second derived group of F_2. For the restrictions to the free group F_3 one has similarly:

$$A^{(2)} \cap F_3 \leqq A^{(3)} \cap F_3 = A \cap F_3.$$

The last term is the second derived group of the group F_3; the first is, by 2.06, the extension of the two-variable second derived group to the group F_3. To show that the inclusion here is, in fact, proper, one has to construct a group generated by three elements, which is not metabelian, while any two of its elements generate a metabelian subgroup. An example of such a group exists[4]); therefore

$$A^{(2)} \cap F_3 < A^{(3)} \cap F_3, \quad \text{hence} \quad A^{(2)} < A^{(3)},$$

the latter by 2.05 and 2.06.

One more application of 2.10 shows that the restriction of the terms of the chain to the free group F_4 are:

$$A^{(2)} \cap F_4 < A^{(3)} \cap F_4 \leqq A^{(4)} \cap F_4 = A \cap F_4.$$

[3]) Neither the notation F^2, nor D_2 or D^2 is convenient in this context.

[4]) The construction of such groups is the subject of a note by B. H. NEUMANN to be published in Proc. Glasgow Math. Ass. 3 (1956).

As before, also the second inclusion is proper, because there exists a non-metabelian group generated by four elements, all whose three-generator subgroups are metabelian[4]). Hence again:

$$A^{(3)} \cap F_4 \subset A^{(4)} \cap F_4, \quad \text{and so} \quad A^{(3)} \subset A^{(4)}.$$

We sum up the situation in this example:

2.12. *If A is the second derived group of F, then*

$$\{0\} = A^{(1)} \subset A^{(2)} \subset A^{(3)} \subset A^{(4)} = A, \quad \text{and so } k = 4.$$

That is: the two-variable metabelian laws form a proper subset of the three-variable metabelian laws, and the three-variable metabelian laws form a proper subset of all metabelian laws.

Since the three corresponding fully invariant subgroups of F all have the same intersection with the group F_2, we see incidentally:

2.13. *The free metabelian group of rank 2 is the free group of rank 2 of at least three different varieties of groups.*

We return once more to the finite basis problem. If there are fully invariant subgroups of F which are infinitely based, then a group like $A^{(2)}$, representing the two-variable metabelian laws, might be thought to provide an example: it is defined as the closure in F of all second commutators of the form $[[x(e_1, e_2), y(e_1, e_2)], [z(e_1, e_2), t(e_1, e_2)]]$. However, GRAHAM HIGMAN has shown[5]), that all two-variable metabelian laws follow from the single law $[[x, y], [-x, y]]$, and so the group $A^{(2)}$ is finitely based.

3. The descending chain determined by a variety

We consider again a fixed variety, given by the fully invariant subgroup V of F. Its free groups of finite rank are $G_n \cong F_n/V_n$, where V_n represents all laws of the variety which involve at most n variables. This time we ask for the variety defined by G_n in the sense of 1.05, that is we want to obtain the set of all laws that hold in G_n, as well as the sets of all laws in m variables that hold in G_n. By 1.06, we obtain all the laws of G_n by representing G_n as a factor group of F. We first note:

3.01. *The set of all the laws of $G_n \cong F_n/V_n$ includes the laws given by V.*

For $F_n/V_n = F_n/(F_n \cap V) \cong \{F_n, V\}/V$; this represents F_n/V_n as a subgroup of F/V, and all the laws of a group hold a fortiori in every subgroup of that group.

To represent F_n/V_n as a factor group of F/V, we use the retraction ε_n of F onto F_n, whose kernel is F_n^* (cf. § 1). As V_n is normal (even fully invariant) in $F_n \cong F/F_n^*$, the group $\{V_n, F_n^*\}$ is normal in F and contains F_n^*. Therefore:

3.02. $$F_n/V_n \cong F/\{V_n, F_n^*\},$$

where the isomorphism is that naturally induced by the mapping ε_n of F onto F_n.

[5]) Unpublished.

In particular:

3.03. *An element $w(e_1, \dots, e_n)$ of F belongs to $\{V_n, F_n^*\}$, if and only if, qua element of F_n it belongs to V_n.*

From 3.01 and 3.02 we have: $V \subset \{V_n, F_n^*\}$, hence $\{V, F_n^*\} \subseteq \{V_n, F_n^*\}$. Since the reverse inclusion holds trivially, we obtain

3.04. *For every n, $\{V_n, F_n^*\} = \{V, F_n^*\}$, and $F_n/V_n \cong F/\{V, F_n^*\}$.*

Now the set of all the laws that hold in F_n/V_n is given by the greatest fully invariant subgroup of F contained in $\{V, F_n^*\}$.

Notation. The greatest fully invariant subgroup of F contained in $\{V, F_n^*\}$ is denoted by W^n.

Then $V \subseteq W^n$; moreover, as $F_n^* > F_{n+1}^*$, also $\{V, F_n^*\} \supseteq \{V, F_{n+1}^*\}$, and so $W^n \supseteq W^{n+1}$. We now have:

3.05. Theorem. *If W^n is the fully invariant subgroup of F representing the set of all laws that hold in the free group of rank n of the variety determined by V, then the groups W^n for $n = 1, 2, \dots$ form a descending chain, whose intersection is V,*

$$W^1 \supseteq W^2 \supseteq \cdots \supseteq W^n \supseteq \cdots \supseteq V, \quad \text{and} \quad V = \bigcap_n W^n.$$

Proof. Only the assertion that the chain contracts to V remains to be proved. Since $V \subseteq W^n \subseteq \{V, F_n^*\}$ for every n, it suffices to prove that

$$\bigcap_n \{V, F_n^*\} = V.$$

Trivially, V is contained in the intersection. Now let w be an element of F that does not belong to V. There exists an integer m such that w belongs to the subgroup F_m of F; since $w \notin V$, it follows that $w \notin V_m$, hence by 3.03 and 3.04, $w \notin \{V, F_m^*\}$, and so $w \notin \bigcap_n \{V, F_n^*\}$, as was to be shown.

3.06. Theorem. *The set of all laws in at most m variables that hold in the free group of rank n of the variety determined by V is given by $W^n \cap F_m$,*

$$\begin{aligned} &\textit{where} \quad W^n \cap F_m = V_m \quad \textit{if} \quad m \leq n, \\ &\textit{and} \quad W^n \cap F_m \supseteq V_m \quad \textit{if} \quad m > n. \end{aligned}$$

Proof. By 3.03, $W^n \cap F_n = V_n$; hence for $m \leq n$:

$$W^n \cap F_m = W^n \cap F_n \cap F_m = V_n \cap F_m = V_m.$$

If $m > n$, $W^n \supseteq W^m$, hence $W^n \cap F_m \supseteq W^m \cap F_m = V_m$.

It is seen again from the case $n = 1$, that some of the inclusions in Theorem 3.05, and therefore also in Theorem 3.06, may be proper: W^1 always contains the whole derived group of F, and its intersections with the groups F_m therefore always contain the whole derived group of the latter. Again, we shall mention a less trivial example at the end of this paragraph.

Combining 3.05 and 3.06 with the results 2.08 and 2.10 of the last paragraph, we now find:

3.07. Theorem. *If G_n is a reduced free group of finite rank n, $G_n \cong F_n/V_n$, then G_n is the free group of rank n of precisely all those varieties determined by fully invariant subgroups U of F which satisfy $V^{(n)} \subseteq U \subseteq W^n$. In particular, the*

varieties determined by $V^{(n)}$ and W^n are the unique greatest and least varieties respectively of which G_n is the free group of rank n.

Proof. As $V^{(n)} \cap F_n = W^n \cap F_n = V_n$, G_n is the free group of rank n of the varieties corresponding to $V^{(n)}$ and W^n, and therefore of every variety between the two.

Conversely, if the fully invariant subgroup U of F determines a variety of which G_n is the free group of rank n, then $U \cap F_n = V_n$, that is V_n is contained in U, and so the closure $V^{(n)}$ of V_n also is contained in U. On the other hand, G_n is a group of the variety determined by U, hence all the laws represented by U hold in G_n, and so U is contained in the fully invariant subgroup representing all the laws of G_n, that is in W^n. This completes the proof.

A different characterization of the terms of the descending chain will show the "duality" between the ascending and the decending chains associated with the fully invariant subgroup V of F.

3.08. Theorem. *The nth term W^n of the descending chain associated with V represents the set of all those laws that hold in every group of the variety determined by V which can be generated by n elements.*

Proof. For the proof we denote by U the set of all the laws holding in every n-generator group of the variety determined by V. Since every n-generator group of this variety is a homomorphic map of F_n/V_n, all the laws of F_n/V_n hold in every one of them; hence $W^n \subseteqq U$. Conversely, F_n/V_n itself is an n-generator group of the variety, and the set of all its laws is precisely W^n; hence $U \subseteqq W^n$, and so $U = W^n$ follows.

Theorem 3.08 is really no more than a restatement of the definition of W^n. The following characterization of W^n is related, but more useful:

3.09. Theorem. *If H is any group defining the variety determined by V, that is, V represents all the laws holding in H, then W^n represents the set of all those laws that hold in every subgroup of H which can be generated by n elements.*

Proof. Again we denote by U the set of all laws holding in every n-generator subgroup of H. With H also every subgroup of H is a group of the variety; 3.08 shows therefore that the laws given by W^n hold in every n-generator subgroup of H, hence $W^n \subseteqq U$. For the converse, we represent H in the form $H \cong F/R$; then V is the greatest fully invariant subgroup of F contained in R. The subgroup of F/R generated by the n elements $e_1 + R, \ldots, e_n + R$ is $\{F_n, R\}/R \cong F_n/R_n$, where $R_n = F_n \cap R$, and V_n is the greatest fully invariant subgroup of F_n contained in R_n. We show that the sets of all the laws of F_n/V_n and of F_n/R_n are the same. Since F_n/R_n is a homomorphic map of F_n/V_n, every law of the latter holds in the former group. But on the other hand, F_n/V_n can be represented as subgroup of a direct power of F_n/R_n (cf. [11]), and every law of a group also holds in every direct power of this group, and therefore also in every subgroup of its direct powers. Thus the laws of F_n/R_n hold in F_n/V_n, and so the two groups possess the same set of laws. But this set of laws is given by W^n, therefore $U \subseteqq W^n$, since the laws given by U hold in every n-generator subgroup of H. Again $U = W^n$ now follows.

We now obtain the counterpart of Theorem 2.08:

3.10. Theorem. *The descending chain determined by the fully invariant subgroup V of F is finite, that is it becomes stationary after a finite number of steps, if and only if, V represents the set of all laws of a finitely generated group.*

If k^* is the least integer such that V represents the set of all laws of a group that can be generated by k^* elements, then

$$W^{k^*-1} > W^{k^*} = W^{k^*+1} = \cdots = V.$$

Proof. If $W^n = V$ for some n, then by definition of W^n, V represents the set of all laws of F_n/V_n, which is generated by n elements. Conversely, if V is the set of all laws of a finitely generated group H, generated by n elements (say), then Theorem 3.08 shows that W^n represents the set of all laws holding in every n-generator group of the variety determined by V, that is, the laws of W^n hold in particular in the group H itself. Therefore $W^n \leqq V$, and so $W^n = V$. The remaining statement of the theorem now follows immediately.

Again it is not known whether every variety can be defined, according to 1.05, by a finitely generated group. Nor is it known whether one of the two chains being finite implies that the other chain also must be finite. If this is true at all, then the two chains certainly need not have the same length. This is shown by the example of the BURNSIDE variety of exponent three, which we discuss as illustration. The facts are, of course, not new: they are due to LEVI and VAN DER WAERDEN ([8]; cf. also [14]), whose results we quote.

3.11. Example. We take V as the fully invariant closure in F of the single word $3e_1$. Thus the ascending chain consists of one term only; $V^{(1)} = V$, $k = 1$.

Next we consider the free groups of rank 1, 2, and 3 of the variety; these are respectively the cycle of order three, and the BURNSIDE groups $B_{2,3} \cong F_2/V_2$ and $B_{3,3} \cong F_3/V_3$. Thus W^1 determines the variety of all abelian groups of exponent three. Further, in both, $B_{2,3}$ and $B_{3,3}$, the law $[x, y, y] = 0$ holds; hence $[e_1, e_2, e_2]$ belongs to V_2, and therefore to every V_n and to V. But $B_{2,3}$ also satisfies the law $[x, y, z] = 0$, while $B_{3,3}$ does not; hence $[e_1, e_2, e_3]$ belongs to W^2, but not to W^3. Finally, all the laws of $B_{3,3}$ also hold in every group $B_{n,3} \cong F_n/V_n$ for $n \geq 3$; hence

$$W^1 > W^2 > W^3 = V, \qquad \text{and so } k^* = 3.$$

It is perhaps not superfluous to show, following the method of this paragraph, how the word $[e_1, e_2, e_3]$, which does not belong to V_3 (and therefore not to W^3, by 3.03), comes to belong to W^2. We recall that W^2 is the greatest fully invariant subgroup of F contained in $\{V, F_2^*\} = \{V_2, F_2^*\}$. Writing x^y for $-y + x + y$, we have:

$$[e_1, e_2, e_3] = -e_3^{[e_1, e_2]} + e_3 \in F_2^* \leqq \{V, F_2^*\}.$$

To show that $[e_1, e_2, e_3]$ belongs to W^2, one has to confirm that every substitution of words x, y, z of F for e_1, e_2, e_3 results in an element of $\{V_2, F_2^*\}$. Now $[x, y, z]$ may be expanded into a sum of commutators of the form $[e_i, e_j, e_k]$

and their transforms. But every such commutator—and therefore also its transforms—belongs to $\{V_2, F_2^*\}$: namely to F_2^*, if all i, j, k are greater than 2; and to V_2, if all i, j, k are equal to 1 or 2. For the remaining cases, $[e_1, e_2, e_k]$ and $[e_1, e_j, e_k]$ with $j, k > 2$ are typical. But

$$[e_1, e_2, e_k] = -e_k^{[e_1, e_2]} + e_k, \quad \text{and} \quad [e_1, e_j, e_k] = [-e_j^{e_1} + e_j, e_k];$$

hence both belong to F_2^*, and therefore to $\{V_2, F_2^*\}$.

Part II

4. The near-ring of endomorphisms of a reduced free group

We consider a fixed reduced free group $G_n \cong F_n/V_n$. At present the rank n may be any finite or infinite cardinal, although we shall restrict ourselves later to finite or countably infinite rank. A basis of G_n, that is, a set of free generators $e_1, e_2, \ldots, e_n$ is chosen, and will remain fixed. Its characteristic property is that every mapping of these generators into the group can be continued to an endomorphism of the group. More precisely:

Every element $g \in G_n$ is representable as a word in the generators $e_1, e_2, \ldots$, $g = g(e_1, \ldots, e_n) = g(\mathfrak{e})$, where $\mathfrak{e}$ stands for the ordered set $e_1, e_2, \ldots, e_n$. This representation is not unique; two words represent the same element of G_n if, and only if, interpreted as elements of the absolutely free group F_n on the same generators e_i, they differ by an element of V_n. Now an endomorphism α of G_n is determined by the images $a_i(\mathfrak{e}) \in G_n$ of the basis elements e_i; the a_i may be arbitrarily chosen, the map of any element g of G_n is then determined:

4.01. *If α is given by $e_i\alpha = a_i(\mathfrak{e})$, $a_i(\mathfrak{e}) \in G_n$ for all i, then*

$$g(e_1, e_2, \ldots, e_n)\,\alpha = g(e_1\alpha, e_2\alpha, \ldots, e_n\alpha) = g(a_1, a_2, \ldots, a_n),$$

or briefly: $g(\mathfrak{e})\,\alpha = g(\mathfrak{a})$.

Because V_n is fully invariant in F_n, only the particular word $g(\mathfrak{a})$, not the element it represents in G_n, depends on the words chosen to represent g and the elements a_i. If now we interpret the ordered set of elements a_i as a vector whose n components are the elements a_i of G_n, then:

4.02. *The set of all endomorphisms α of G_n is represented by the set of all vectors with n components a_i in G_n, the representation being given by $e_i\alpha = a_i$ for all i.*

Two vectors represent the same endomorphism if, and only if, corresponding components, interpreted as elements of F_n, are congruent modulo V_n.

Accordingly, we may now always interpret a small Greek letter as signifying either an endomorphism of G_n or an n-vector with components in G_n. This dual interpretation will prove useful.

4.03. Definition. *If α and β are endomorphisms of G_n, then the the sum and the product of α and β are, respectively, the endomorphisms given by:*

$$e_i(\alpha + \beta) = e_i\alpha + e_i\beta \quad \text{and} \quad e_i(\alpha\beta) = (e_i\alpha)\,\beta, \quad \text{for all } i.$$

Note that, because α and β are homomorphisms, the product of two endomorphisms is defined in agreement with the customary definition of the product of two mappings. But not so the sum: The sum of two mappings μ, ν of a group H into itself is usually defined as the mapping given by $h(\mu+\nu)=h\mu+h\nu$ for *all* $h \in H$, and is therefore not in general an endomorphism of H, even if μ and ν both are endomorphisms. In our case, $\alpha+\beta$ is by definition an endomorphism; but 4.01 shows that the map of an arbitrary element g under $\alpha+\beta$ differs in general from $g\alpha+g\beta$, as G_n is not in general abelian.

The algebraic system of all endomorphisms of G_n with addition and multiplication defined by 4.03 will be denoted by Γ_n. Expressing the algebraic operations in Γ_n in terms of the vector representation, we have:

4.04. *If α has components a_i, and β has components b_i, then $\alpha+\beta$ and $\alpha\beta$ are given by*

$$e_i(\alpha+\beta) = a_i(\mathfrak{e}) + b_i(\mathfrak{e}) \qquad \text{for all } i,$$

and

$$e_i(\alpha\beta) = a_i(\mathfrak{b}) = a_i(b_1, \ldots, b_n) \qquad \text{for all } i.$$

Thus, the sum of the endomorphisms is represented by the sum of the vectors; the product is represented by the map of the vector α under the endomorphism β. Or, expressed differently, a product $\alpha\beta$ may be read either as a product of two endomorphism, or as the map of the vector α under the endomorphism β, whichever is more useful in the context. This shows almost immediately:

4.05. Theorem. *Γ_n is a near-ring. More precisely: 1. Γ_n is a semigroup under multiplication, with an identity element ε given by $e_i\varepsilon = e_i$ for all i. 2. Γ_n is a group with respect to addition; the additive group is the cartesian (or unrestricted direct) sum of n isomorphic copies of G_n. The zero of addition, also denoted by 0, is a two-sided zero of multiplication. 3. Γ_n satisfies the right distributive law, $(\alpha+\beta)\gamma = \alpha\gamma+\beta\gamma$ for all $\alpha, \beta, \gamma \in \Gamma_n$.*

Proof. The first statement follows from the interpretation of $\alpha\beta$ as a product of endomorphisms; the second from the interpretation of $\alpha+\beta$ as a sum of two n-vectors with components in G_n. The distributive law is confirmed directly from 4.03:

$$\begin{aligned} e_i(\alpha+\beta)\gamma = (e_i\alpha + e_i\beta)\gamma &= (e_i\alpha)\gamma + (e_i\beta)\gamma, && \text{as } \gamma \text{ is homomorphic,} \\ &= e_i(\alpha\gamma+\beta\gamma), && \text{for all } i. \end{aligned}$$

The endomorphism 0 maps every element, hence every vector, onto 0; therefore $\alpha 0 = 0$ for all α. But $0\alpha = 0$ follows from the distributive law, so that 0 is, in fact, the two-sided zero of multiplication.

The multiplicative semigroup, and the additive group, of Γ_n will be denoted by $\Gamma_n(\cdot)$ and $\Gamma_n(+)$ respectively. The kth summand of $\Gamma_n(+)$ in its representation as cartesian sum consists of all endomorphisms α^k,

4.06 $\qquad e_i\alpha^k = 0 \quad \text{for} \quad i \neq k, \quad e_k\alpha_k = a_k(\mathfrak{e}) \quad \text{with} \quad a_k(\mathfrak{e}) \in G_n.$

If in particular we denote by ε_j^k the endomorphism given by

4.07 $\qquad e_i\varepsilon_j^k = 0 \quad \text{for} \quad i \neq k, \quad e_k\varepsilon_j^k = e_j,$

then every element α^k is a finite (non-commutative) sum of elements ε_j^k,

$$\alpha^k = \sum_j n_j \varepsilon_j^k, \quad \text{where the } n_j \text{ are integers.}$$

Hence we have from the structure of $\Gamma_n(+)$:

4.08. Theorem. *If n is finite, every element of Γ_n is a finite sum of integral multiples of the n^2 elements ε_j^k.*

The significance of the theorem stems from the following fact:

4.09. Theorem. *In every near-ring Γ_n, every element ε_j^k distributes also as left-hand factor,*

$$\varepsilon_j^k(\alpha+\beta) = \varepsilon_j^k\alpha + \varepsilon_j^k\beta \quad \text{for all } j, k \text{ and all} \quad \alpha, \beta \in \Gamma_n.$$

Proof. As ε_j^k maps the ith generator onto zero, whenever $i \neq k$, and zero is mapped onto itself by every endomorphism, it follows that the ith component of both, the right-hand and the left-hand side, is zero. For the kth components we have

$$e_k\varepsilon_j^k(\alpha+\beta) = e_j(\alpha+\beta) = e_j\alpha + e_j\beta = e_k\varepsilon_j^k\alpha + e_k\varepsilon_j^k\beta = e_k(\varepsilon_j^k\alpha + \varepsilon_j^k\beta).$$

The preceding remarks show that we may think of Γ_n as a generalized matrix ring. The "matrix" representing α is given by:

$$\alpha \leftrightarrow [a_i(e_j)] = \begin{pmatrix} a_1(e_1, e_2, \ldots, e_n) \\ \ldots\ldots\ldots\ldots \\ a_n(e_1, e_2, \ldots, e_n) \end{pmatrix};$$

that is, the "rows" correspond to the components of α, the "columns" to the ordered set of generators e_i. If n is infinite, the "matrices" still are row-finite. The multiplication rule is given by

$$\alpha\beta \leftrightarrow [a_i(e_j)]\,[b_j(e_k)] = [a_i(b_j(e_k))].$$

If G_n is a free abelian group, this interpretation gives precisely the representation of the ring of endomorphisms as a ring of matrices over the integers (or over the integers modulo m, if G_n is the free abelian group of exponent m).

We conclude this paragraph with a discussion of units and right regular elements of $\Gamma_n(\cdot)$. We need a short digression: Deriving from a well known question posed by H. HOPF (cf. e.g. [12]), we call a group H a HOPF-group if it is not isomorphic to a proper factor group of itself. A reduced free group G_n can be a HOPF-group only if n is finite. Whether this condition is sufficient, is an unsolved problem; so far, only certain classes of reduced free groups of finite rank have been shown to be HOPF-groups (cf. e.g. [1], [5], [10]). From the definition we have:

4.10. *A reduced free group G_n of finite rank is a* HOPF*-group if, and only if, every n elements that generate G_n generate it freely.*

With this situation in mind we define:

4.11. Definition. *An element α is a unit of Γ_n, if it possesses a unique (right- and left-)inverse α^{-1} in Γ_n.*

An element α of Γ_n is right regular, if $\zeta\alpha=\eta\alpha$ implies $\zeta=\eta$; or, equivalently, if $\xi\alpha=0$ implies $\xi=0$.

Clearly, the units of Γ_n are exactly the automorphisms of G_n. Every unit is right regular, and the right regular elements of Γ_n are exactly the iso-endomorphisms of G_n, that is the isomorphic mappings of G_n into itself.

4.12. Theorem: *The relation $\alpha\beta=\varepsilon$ in Γ_n implies that α, or equivalently β, is a unit (and then $\beta=\alpha^{-1}$) if, and only if, the group G_n is a* HOPF*-group.*

Proof. Assume first that G_n is a HOPF-group, and that $\alpha\beta=\varepsilon$ in Γ_n. If α and β have components a_i and b_i respectively, we have therefore $a_i(b_1,\ldots,b_n)=e_i$ for $i=1,\ldots,n$. Hence the elements $b_1,\ldots,b_n$ generate G_n, and therefore generate it freely. This means that the mapping β given by $e_i\beta=b_i$ $(i=1,\ldots,n)$ is an automorphism of G_n, and β^{-1} exists. But then $(\alpha\beta)\beta^{-1}=\beta^{-1}=\alpha(\beta\beta^{-1})=\alpha$, and so α also is a unit.

Next assume that G_n is not a HOPF-group, that is, it possesses a proper factor group isomorphic to itself. The natural homomorphism onto this factor group followed by the isomorphism of the factor group onto G_n constitutes a proper endomorphism β of G_n. Let $e_i\beta=b_i$ $(i=1,\ldots,n)$. Since β is a mapping onto G_n, there exist elements a_i in G_n such that $a_i\beta=e_i$ $(i=1,\ldots,n)$. We define α by $e_i\alpha=a_i$ $(i=1,\ldots,n)$. Then $e_i\alpha\beta=e_i$ for all i, hence $\alpha\beta=\varepsilon$. But β is a proper endomorphism and so possesses no inverse, which proves 4.12.

We note that the α of the proof is right regular: $\xi\alpha=0$ implies $(\xi\alpha)\beta=0=\xi(\alpha\beta)=\xi$. Thus α is an iso-endomorphism, and our proof gives as corollary:

4.13. BAER'S Theorem ([1]). *If a reduced free group is isomorphic to a proper factor group, then also to a proper subgroup.*

5. The near-ring Γ_ω

We now assume n to be finite or countably infinite, so that the near-rings are those corresponding to reduced free groups of finite or countably infinite rank. The latter will now be denoted again by G, without subscript. However, we restrict ourselves even further: since $\Gamma_{\aleph_0}(+)$ is the unrestricted direct sum of $\aleph_0$ copies of G, Theorem 4.08 does not hold. But the existence of a "distributive basis" as given by 4.08 and 4.09 is essential for the ideal theory, which we wish to use. We therefore consider only a sub-near-ring Γ_ω of $\Gamma_{\aleph_0}$, namely—in terms of the generalized matrix representation—the near-ring of all row- and column-finite matrices of $\Gamma_{\aleph_0}$.

For every finite n we consider the subset of $\Gamma_{\aleph_0}$ consisting of all those endomorphisms α whose first n components involve only the generators $e_1,\ldots,e_n$, and whose remaining components are zero:

$$e_i\alpha=a_i(e_1,\ldots,e_n) \quad \text{for} \quad i=1,\ldots,n; \qquad e_i\alpha=0 \quad \text{for} \quad i>n.$$

Evidently these elements form a near-ring contained in $\Gamma_{\aleph_0}$ which is isomorphic to Γ_n. We denote it by the the same symbol Γ_n without danger of

confusion. Then $\Gamma_1 \subset \Gamma_2 \subset \cdots \subset \Gamma_n \subset \cdots$, and we put

$$\Gamma_\omega = \bigcup_n \Gamma_n. \tag{5.01}$$

Γ_ω is clearly a near-ring; it consists of all those endomorphisms of G which map the subgroup generated by $e_1, \ldots, e_n$ (for some n) into itself, and the complimentary subgroup on zero.

The additive group $\Gamma_\omega(+)$ of Γ_ω is the (restricted) direct sum of countably many isomorphic copies of G. Hence the essential basis property holds; we summarise it for the case of Γ_ω:

5.02. Theorem. *If ε_j^k is given by $e_i \varepsilon_j^k = 0$ for $i \neq k$, and $e_k \varepsilon_j^k = e_j$, then*
(i) *every element $\xi \in \Gamma_\omega$ is a finite sum of integral multiples of elements ε_j^k,*

$$\xi = \sum_{j,k} n_{j,k} \varepsilon_j^k \qquad (n_{j,k} \text{ integers});$$

(ii) *every ε_j^k distributes also as left-hand factor,*

$$\varepsilon_j^k(\alpha + \beta) = \varepsilon_j^k \alpha + \varepsilon_j^k \beta \quad \text{for all } j, k \text{ and all } \alpha, \beta \in \Gamma_\omega.$$

The near-ring Γ_ω possesses no identity element, nor does it possess units or right-regular elements. We call the identity elements of its sub-near-rings Γ_n "relative identities", and denote them by ε_n. The ε_n are just the retractive endomorphisms considered in the case of the absolutely free groups in § 1. As then, they can again be used to pick out mechanically the near-rings Γ_n from the whole of Γ_ω:

5.03. *An element ξ of Γ_ω belongs to Γ_n if, and only if, $\varepsilon_n \xi \varepsilon_n = \xi$. For any element η of Γ_ω, $\varepsilon_n \eta \varepsilon_n$ belongs to Γ_n, and $\Gamma_n = \varepsilon_n \Gamma_\omega \varepsilon_n$.*

Proof. If $\xi \in \Gamma_n \subset \Gamma_\omega$, then $\varepsilon_n \xi \varepsilon_n = \xi$, as ε_n is the identity of Γ_n. Conversely, if $\xi = \varepsilon_n \xi \varepsilon_n$, then only the first n components of ξ are non-trivial, because the first factor ε_n has this property; and ξ depends on $e_1, \ldots, e_n$ only, because the last factor ε_n has this property.

If $\eta \in \Gamma_\omega$, then $\varepsilon_n \eta \varepsilon_n = \varepsilon_n^2 \eta \varepsilon_n^2$, hence $\varepsilon_n \eta \varepsilon_n \in \Gamma_n$ by the first part. Therefore $\varepsilon_n \Gamma_\omega \varepsilon_n \subseteq \Gamma_n$. But also $\Gamma_n = \varepsilon_n \Gamma_n \varepsilon_n \subseteq \varepsilon_n \Gamma_\omega \varepsilon_n$, hence $\Gamma_n = \varepsilon_n \Gamma_\omega \varepsilon_n$.

Further, one confirms easily:

5.04. *Every ε_n distributes also as left-hand factor: $\varepsilon_n(\alpha + \beta) = \varepsilon_n \alpha + \varepsilon_n \beta$, for all n, and all $\alpha, \beta \in \Gamma_\omega$.*

6. Ideals and homomorphisms of Γ_n and Γ_ω

The relation between ideals and homomorphisms in the near-rings Γ_n and Γ_ω relies entirely on the existence of an additive basis which consists of elements distributing on both sides. Beyond that the fact that these near-rings are in character near-rings of matrices over a group is used essentially. Most statements in this and the following paragraph hold equally in Γ_n and Γ_ω; wherever this is the case we simply use Γ.

6.01. Definition. *The subset $\mathfrak{A}$ of Γ is a left ideal, right ideal, or (two-sided) ideal, if $\mathfrak{A}$ is a normal subgroup of $\Gamma(+)$, and $\xi\mathfrak{A} \subseteq \mathfrak{A}$, $\mathfrak{A}\eta \subseteq \mathfrak{A}$, $\xi\mathfrak{A}\eta \subseteq \mathfrak{A}$ respectively for all $\xi, \eta \in \Gamma$.*

6.02. *For every ideal $\mathfrak{A}$ of Γ, the relation "$\beta \equiv \gamma \bmod \mathfrak{A}$ if, and only if, $\beta - \gamma \in \mathfrak{A}$" is a congruence relation on the near-ring Γ.*

Proof. Since $\mathfrak{A}$ is, by definition, a normal subgroup of $\Gamma(+)$, we need only prove that, if $\beta \equiv \gamma \bmod \mathfrak{A}$ and $\beta' \equiv \gamma' \bmod \mathfrak{A}$, then $\beta\beta' \equiv \gamma\gamma' \bmod \mathfrak{A}$. From the premise we have:

$$\beta = \gamma + \alpha, \quad \beta' = \gamma' + \alpha', \quad \text{where} \quad \alpha, \alpha' \in \mathfrak{A}.$$

Hence

$$\beta\beta' = (\gamma + \alpha)(\gamma' + \alpha') = \gamma(\gamma' + \alpha') + \alpha(\gamma' + \alpha') = \gamma(\gamma' + \alpha') + \alpha'',$$

where $\alpha'' \in \mathfrak{A}$, as $\mathfrak{A}$ is a right ideal. To prove that

$$\gamma(\gamma' + \alpha') = \gamma\gamma' + \alpha^* \qquad \text{with } \alpha^* \in \mathfrak{A},$$

we note that (i) this is true if $\gamma = \varepsilon_j^k$; for then we have

$$\varepsilon_j^k(\gamma' + \alpha') = \varepsilon_j^k\gamma' + \varepsilon_j^k\alpha',$$

where the last term belongs to $\mathfrak{A}$, since $\mathfrak{A}$ is a left ideal;

(ii) if $\gamma_1(\gamma' + \alpha') = \gamma_1\gamma' + \alpha_1$ and $\gamma_2(\gamma' + \alpha') = \gamma_2\gamma' + \alpha_2$ with $\alpha_1, \alpha_2 \in \mathfrak{A}$, then

$$\begin{aligned}(\gamma_1 + \gamma_2)(\gamma' + \alpha') &= \gamma_1(\gamma' + \alpha') + \gamma_2(\gamma' + \alpha') = \gamma_1\gamma' + \alpha_1 + \gamma_2\gamma' + \alpha_2 \\ &= \gamma_1\gamma' + \gamma_2\gamma' - \gamma_2\gamma' + \alpha_1 + \gamma_2\gamma' + \alpha_2 = (\gamma_1 + \gamma_2)\gamma' + \alpha_3,\end{aligned}$$

where α_3 belongs to $\mathfrak{A}$, as $\mathfrak{A}$ is a normal subgroup of $\Gamma(+)$. By 4.08 and 5.02 every element γ is a finite sum of elements ε_j^k; therefore (i) and repeated application of (ii) prove 6.02.

It now follows in the usual way that addition and multiplication of congruence classes can be defined in the natural manner, and that the congruence classes then form a near-ring, the quotient near-ring $\Gamma/\mathfrak{A}$. The mapping which maps every element of Γ on its class modulo $\mathfrak{A}$ is a homomorphism of Γ onto $\Gamma/\mathfrak{A}$, of which $\mathfrak{A}$ is the kernel. Conversely, if $\hat{\varrho}$ is any homomorphic mapping of Γ onto a near-ring $\Gamma\hat{\varrho} = \hat{\Gamma}$, then the kernel of the mapping is an ideal $\mathfrak{A}$ of Γ. For it is clearly a normal subgroup of $\Gamma(+)$; and if $\alpha \in \mathfrak{A}$, that is $\alpha\hat{\varrho} = 0$, then $(\xi\alpha\eta)\hat{\varrho} = \xi\hat{\varrho} \cdot 0 \cdot \eta\hat{\varrho} = 0$, hence $\xi\alpha\eta \in \mathfrak{A}$ for all $\xi, \eta \in \Gamma$. Therefore we have:

6.03. Theorem. *In every near-ring Γ_n or Γ_ω the ideals are precisely the kernels of the homomorphisms of Γ_n or Γ_ω onto a near-ring $\Gamma_n\hat{\varrho}$ or $\Gamma_\omega\hat{\varrho}$ respectively.*

In the usual way one deduces further:

6.04. Theorem. *The map of an ideal (left ideal, right ideal) of Γ under the homomorphism $\hat{\varrho}$ is an ideal (left ideal, right ideal) of $\Gamma\hat{\varrho}$. The ideals (left ideals, right ideals) of $\Gamma\hat{\varrho}$ are in one-one correspondence to those ideals (left ideals, right ideals) of Γ which contain the kernel $\mathfrak{A}$ of $\hat{\varrho}$.*

We turn to an alternative characterization of the ideals (left ideals, right ideals) in terms of the components of their elements. We recall, that the free group associated with Γ is G_n, if $\Gamma = \Gamma_n$, and G, if $\Gamma = \Gamma_\omega$.

6.05. Theorem. *The subset $\mathfrak{A}$ of Γ is a left ideal if, and only if, the components of the elements of $\mathfrak{A}$ vary independently over the elements of a normal subgroup N of the free group associated with Γ.*

The subset $\mathfrak{A}$ of Γ is a right ideal if, and only if, it is a normal subgroup of $\Gamma(+)$, and for each fixed i, the ith components of the elements of $\mathfrak{A}$ form a fully invariant subgroup $A(i)$ of the free group associated with Γ.

Corollary. *The subset $\mathfrak{A}$ of Γ is a right ideal if the components of the elements of $\mathfrak{A}$ vary independently over a set of fully invariant subgroups $A(i)$ of the free group associated with Γ.*

Proof. The argument is almost the same in the two cases $\Gamma = \Gamma_n$ and $\Gamma = \Gamma_\omega$.

1. If $\mathfrak{A}$ is a left ideal, it is a normal subgroup of $\Gamma(+)$, hence the set of all i^{th} components of elements of $\mathfrak{A}$ is a normal subgroup $N(i)$ of G_n or G respectively. If $a_i \in N(i)$, there exists an element $\alpha \in G$ such that $e_i \alpha = a_i$. Since $\mathfrak{A}$ is a left ideal, also $\varepsilon_i^k \alpha \in \mathfrak{A}$. But $e_k \varepsilon_i^k \alpha = e_i \alpha = a_i$, hence $a_i \in N(k)$. It follows that for all i and k, $N(i) = N(k) = N$. Now let $a_1, \ldots, a_n$[6]) be any given elements of N. Then a_i occurs as first component of some element α_i of $\mathfrak{A}$. With α_i also $\varepsilon_1^i \alpha_i$ belongs to $\mathfrak{A}$, and

$$e_i \varepsilon_1^i \alpha_i = e_1 \alpha_i = a_i, \qquad e_j \varepsilon_1^i \alpha_i = 0 \quad \text{for } j \neq i. \tag{6.06}$$

But then the sum $\sum_{i=1}^{n} \varepsilon_1^i \alpha_i$ also belongs to $\mathfrak{A}$, and its components are the prescribed elements a_i of N for $i = 1, \ldots, n$, and the remaining components (in the case $\Gamma = \Gamma_\omega$) are zero.

Conversely, if $\mathfrak{A}$ satisfies the premise of the first part of 6.05, it is evidently a normal subgroup of $\Gamma(+)$. Moreover, if $\alpha \in \mathfrak{A}$, and α has components a_i, and if $\xi \in \Gamma$ is any element with components $x_i(e_1, \ldots, e_n)$[6]), then the components of $\xi \alpha$ are the words $x_i(a_1, \ldots, a_n)$[6]) formed with the elements a_i of N, hence they belong to N, and so $\xi \alpha \in \mathfrak{A}$ follows; that is $\mathfrak{A}$ is a left ideal.

2. If $\mathfrak{A}$ is a right ideal of Γ, it is a normal subgroup of $\Gamma(+)$; hence the elements occuring as i^{th} components of elements of $\mathfrak{A}$ form a normal subgroup $A(i)$ of G_n or G respectively. Moreover, if $\alpha \in \mathfrak{A}$, and its ith component is $a_i(e_1, \ldots, e_n)$[6]), let η be any element of Γ with components $y_1, \ldots, y_n$[6]) (and remaining components zero in the case $\Gamma = \Gamma_\omega$). Then the ith component of $\alpha \eta$ is $a_i(y_1, \ldots, y_n)$. But $\alpha \eta \in \mathfrak{A}$, therefore $a_i(y_1, \ldots, y_n) \in A(i)$, which shows that $A(i)$ is a fully invariant subgroup of G_n or G respectively.

Conversely, if the $A(i)$ are fully invariant in the free group associated with Γ, then one confirms in the same way that $\alpha \eta \in \mathfrak{A}$ for all $\eta \in \Gamma$, provided $\alpha \in \mathfrak{A}$. Since $\mathfrak{A}$ is given to be a normal subgroup of $\Gamma(+)$, it follows that it is a right ideal.

If the components of $\mathfrak{A}$ run independently over the fully invariant, that is certainly normal, subgroups $A(i)$, then $\mathfrak{A}$ is ensured to be normal in $\Gamma(+)$; hence the corollary.

[6]) In the case $\Gamma = \Gamma_n$, the n occurring here is the rank n of G_n in the case $\Gamma = \Gamma_\omega$, n is an arbitrary positive integer.

Combining the two parts of 6.05, we finally have for ideals:

6.07. Theorem. *The ideals $\mathfrak{A}$ of Γ are in one-one correspondence with the fully invariant subgroups A of the free group associated with Γ. The correspondence is given by: $\alpha \in \mathfrak{A}$ if, and only if, $e_i \alpha \in A$ for all i.*

Before we use this correspondence between ideals and fully invariant subgroups to give a description of the homomorphic maps of a near-ring Γ, we note a consequence of the proof of 6.05, which we shall use later. In the case of two-sided ideals, it is sufficient to know that $\mathfrak{A}$ is a subgroup of $\Gamma(+)$ which is closed under left- and right multiplication by elements of Γ; it then follows that $\mathfrak{A}$ is a normal subgroup of $\Gamma(+)$:

6.08. *The subset $\mathfrak{A}$ of Γ is an ideal if (and only if) it is closed under addition, and left- and right multiplication by elements of Γ.*

Proof. From the assumptions it follows that $\mathfrak{A}$ is closed under subtraction, hence is a subgroup of $\Gamma(+)$. We have to show that it is a normal subgroup that is, that $\alpha \in \mathfrak{A}$ and $\eta \in \Gamma$ implies $-\eta+\alpha+\eta \in \mathfrak{A}$.

Let α have components $a_1, \ldots, a_n$, η components $y_1, \ldots, y_n$[6]). The left ideal property implies that any component a_i can be "isolated" in $\mathfrak{A}$ (as in 6.06); that is, there belongs to $\mathfrak{A}$ also the element $\varepsilon_i^i \alpha$, for every i, whose only non-zero component is the ith component a_i of α. If $a_i = a_i(e_1, \ldots, e_n)$[6]), and η_i is the element with components $-y_i+e_j+y_i$ $(j=1, \ldots, n)$, where y_i belongs to G_n or G respectively (and in the case of $\Gamma = \Gamma_\omega$ all further components of η are zero), then the ith component of $\varepsilon_i^i \alpha \eta_i$ is $-y_i+a_i+y_i$, and all its other components are zero. Therefore $-\eta+\alpha+\eta = \sum_{i=1}^{n} \varepsilon_i^i \alpha \eta_i$, which clearly belongs to $\mathfrak{A}$.

The homomorphism of Γ whose kernel is $\mathfrak{A}$ reduces every element of Γ modulo $\mathfrak{A}$, that is it reduces the components of the elements of Γ modulo A. But the factor group of a reduced free group with respect to one of its fully invariant subgroups is itself a reduced free group of the same rank. Therefore the Theorems 6.03 and 6.07 now lead to:

6.09. Theorem. *Every homomorphic map of the near-ring Γ_n or Γ_ω of endomorphisms of the reduced free group G_n or G respectively is itself a near-ring $\hat{\Gamma}_n$ or $\hat{\Gamma}_\omega$. If $\mathfrak{A}$ is the kernel of the homomorphism, A the corresponding fully invariant subgroup of G_n or G respectively, then the reduced free group associated with $\hat{\Gamma}_n$ or $\hat{\Gamma}_\omega$ is the factor group G_n/A or G/A respectively.*

Corollary. *Every near-ring Γ_n or Γ_ω is a homomorphic map of the corresponding near-ring Φ_n or Φ_ω associated with the absolutely free group F_n or F respectively.*

The corollary permits an amusing formulation of HOPF's problem for reduced free groups of finite rank: We have seen (Theorem 4.12) that G_n is a HOPF group if, and only if, $\alpha\beta = \varepsilon$ in Γ_n implies $\beta = \alpha^{-1}$, or equivalently, $\beta\alpha = \varepsilon$. As $\Gamma_n \cong \Phi_n/\mathfrak{A}$, where $\mathfrak{A}$ is the ideal corresponding to the fully invariant subgroup of F_n which defines G_n, we have:

6.10. *Every reduced free group of finite rank is a* HOPF *group (that is, a group not isomorphic to any proper factor group of itself) if, and only if, for every n, and for any two elements α and β of Φ_n, the least ideal containing $\alpha\beta - \varepsilon$ also contains $\beta\alpha - \varepsilon$.*

7. Lattice operations and multiplication of ideals in Γ_n and Γ_ω

Except for the last theorem, this paragraph makes no explicit reference to the underlying free groups; the elementary algebra of ideals depends entirely on the matrix character of the near-rings Γ_n and Γ_ω. Again we write Γ whenever a statement applies equally to Γ_n and Γ_ω.

Immediately from the definition of ideals (6.01) we have:

7.01. *The intersection of two ideals (left ideals, right ideals) of Γ is an ideal (left ideal, right ideal) of Γ.*

7.02. Definition. *The sum $\mathfrak{A}+\mathfrak{B}$ of two ideals (left ideals, right ideals) of Γ is the least ideal (left ideal, right ideal) of Γ containing $\mathfrak{A}$ and $\mathfrak{B}$.*

7.03. *The sum $\mathfrak{A}+\mathfrak{B}$ of two ideals (left ideals, right ideals) consists of all elements of the form $\alpha+\beta$, $\alpha\in\mathfrak{A}$, $\beta\in\mathfrak{B}$.*

Proof. In all three cases, $\mathfrak{A}$ and $\mathfrak{B}$ are normal subgroups of $\Gamma(+)$. Hence the set of all sums $\alpha+\beta$ with $\alpha\in\mathfrak{A}$, $\beta\in\mathfrak{B}$, is the least normal subgroup of $\Gamma(+)$ containing both $\mathfrak{A}$ and $\mathfrak{B}$. That this normal subgroup is closed under right multiplication by elements of Γ, if $\mathfrak{A}$ and $\mathfrak{B}$ are right ideals, follows from the distributive law in Γ. That it is closed under left multiplication by elements of Γ, if $\mathfrak{A}$ and $\mathfrak{B}$ are left ideals, follows from the existence of a distributive basis for Γ. One has

(i) $$\varepsilon_i^k(\alpha+\beta) = \varepsilon_j^k\alpha + \varepsilon_j^k\beta = \alpha' + \beta';$$

and

(ii) if $\xi_1(\alpha+\beta) = \alpha_1+\beta_1$ and $\xi_2(\alpha+\beta) = \alpha_2+\beta_2$,

then

$$(\xi_1+\xi_2)(\alpha+\beta) = \xi_1(\alpha+\beta) + \xi_2(\alpha+\beta) = \alpha_1+\beta_1+\alpha_2+\beta_2$$
$$= \alpha_1+\alpha_2-\alpha_2+\beta_1+\alpha_2+\beta_2 = \alpha_3+\beta_3,$$

because $-\alpha_2+\beta_1+\alpha_2$ belongs to $\mathfrak{B}$, as $\mathfrak{B}$ is normal in $\Gamma(+)$.

(i) and repeated application of (ii) give $\xi(\alpha+\beta)=\alpha''+\beta''$ for all $\xi\in\Gamma$. It is clear from Theorem 6.07 that:

7.04. *The fully invariant subgroups corresponding to the intersection and to the sum of two ideals $\mathfrak{A}$ and $\mathfrak{B}$ are the intersection and the sum respectively of the fully invariant subgroups corresponding to $\mathfrak{A}$ and $\mathfrak{B}$.*

The two-step argument used in 7.03 shows further:

7.05. *The ideal generated by a set of elements α_i of Γ, that is, the least ideal containing the elements α_i, consists of all finite sums of the form $\sum_i \xi_i\alpha_i\eta_i$, $\xi_i, \eta_i\in\Gamma$ (where, of course, the same element α_i may occur repeatedly with different left- and right factors).*

Proof. Every sum $\sum_i \xi_i \alpha_i \eta_i$ clearly belongs to every ideal containing the elements α_i. It therefore remains to show that the set of all such sums is itself an ideal. We use 6.08: The set of all these sums is additively closed; it is closed under right multiplication because of the distributive law; finally $\xi \sum \xi_i \alpha_i \eta_i = \sum \xi_i' \alpha_i \eta_i'$ for every $\xi \in \Gamma$, because (i) this is true when $\xi = \varepsilon_j^k$, as this distributes over the sum; and (ii) if $\xi_1 \sum \xi_i \alpha_i \eta_i = \sum \xi_i' \alpha_i \eta_i'$ and $\xi_2 \sum \xi_i \alpha_i \eta_i = \sum \xi_i'' \alpha_i \eta_i''$, then, using the distributive law,

$$(\xi_1 + \xi_2) \sum \xi_i \alpha_i \eta_i = \sum \xi_i' \alpha_i \eta_i' + \sum \xi_i'' \alpha_i \eta_i'',$$

and this is itself a finite sum of the same form.

Now we define the product of ideals in the usual way:

7.06. Definition. *The product $\mathfrak{A}_l \mathfrak{B}_r$ of a left ideal $\mathfrak{A}_l$ and a right ideal $\mathfrak{B}_r$ of Γ is the ideal generated in Γ by all products $\alpha\beta$, where $\alpha \in \mathfrak{A}_l$ and $\beta \in \mathfrak{B}_r$.*

As $\xi \alpha \beta \eta = (\xi \alpha)(\beta \eta) = \alpha' \beta'$, $\alpha' \in \mathfrak{A}_l$, $\beta' \in \mathfrak{B}_r$, 7.05 gives:

7.07. Theorem. *The product $\mathfrak{A}_l \mathfrak{B}_r$ consists of all finite sums of the form $\sum \alpha_i \beta_i$, where $\alpha_i \in \mathfrak{A}_l$, $\beta_i \in \mathfrak{B}_r$.*

If both factors are two-sided ideals $\mathfrak{A}$, $\mathfrak{B}$, then each term $\alpha_i \beta_i$ belongs to $\mathfrak{A} \cap \mathfrak{B}$, and this is additively closed. Hence:

7.08. *If $\mathfrak{A}$ and $\mathfrak{B}$ are ideals of Γ, then $\mathfrak{A}\mathfrak{B} \subseteq \mathfrak{A} \cap \mathfrak{B}$.*

By way of illustration we mention here the *mutual commutator ideal* $[\mathfrak{A}, \mathfrak{B}]$ of two ideals $\mathfrak{A}$ and $\mathfrak{B}$. It is the ideal corresponding to the mutual commutator group $[A, B]$ of A and B, itself a fully invariant subgroup of the group associated with Γ. The ideal $[\mathfrak{A}, \mathfrak{B}]$ is the product of the ideal corresponding to the derived group, briefly the derived ideal, as left hand factor with the right ideal consisting of all those elements of Γ, whose first components lie in A, whose second components lie in B, and whose further components are zero (cf. the corollary of 6.05). One has:

7.09. $[\mathfrak{A}, \mathfrak{B}] \subseteq \mathfrak{A} \cap \mathfrak{B}$ *for any two ideals $\mathfrak{A}$ and $\mathfrak{B}$ of Γ.*

We put together some of the relations between the lattice operations and ideal multiplication which hold in every near-ring Γ. As usual, the ideal consisting of the zero endomorphism only acts as a zero of multiplication, the whole near-ring Γ acts as the two-sided identity of ideal multiplication. Further:

7.10. $\mathfrak{A}(\mathfrak{B}_1 \cap \mathfrak{B}_2) \subseteq \mathfrak{A}\mathfrak{B}_1 \cap \mathfrak{A}\mathfrak{B}_2$, $(\mathfrak{B}_1 \cap \mathfrak{B}_2)\mathfrak{A} \subseteq \mathfrak{B}_1\mathfrak{A} \cap \mathfrak{B}_2\mathfrak{A}$.
This is an immediate consequence of 7.07.

7.11. $(\mathfrak{B}_1 + \mathfrak{B}_2)\mathfrak{A} = \mathfrak{B}_1\mathfrak{A} + \mathfrak{B}_2\mathfrak{A}$; $\mathfrak{A}\mathfrak{B}_1 + \mathfrak{A}\mathfrak{B}_2 \subseteq \mathfrak{A}(\mathfrak{B}_1 + \mathfrak{B}_2)$.
$\mathfrak{A}(\mathfrak{B}_1 + \mathfrak{B}_2) \subseteq \mathfrak{A}\mathfrak{B}_1 + \mathfrak{A}\mathfrak{B}_2 + \mathfrak{A} \cap [\mathfrak{B}_1, \mathfrak{B}_2] \subseteq \mathfrak{A}\mathfrak{B}_1 + \mathfrak{A}\mathfrak{B}_2 + \mathfrak{A} \cap \mathfrak{B}_1 \cap \mathfrak{B}_2$.

The first line follows from 7.03 and 7.07; the second from the fact that $\alpha(\beta_1 + \beta_2)$ differs from $\alpha\beta_1 + \alpha\beta_2$ by elements in the mutual commutator $[\mathfrak{B}_1, \mathfrak{B}_2]$ which must also belong to $\mathfrak{A}$, as $\mathfrak{A}(\mathfrak{B}_1 + \mathfrak{B}_2) \subseteq \mathfrak{A}$.

We note as a corollary:

7.12. *If* $\mathfrak{A}+\mathfrak{B}=\Gamma$, *then* $\mathfrak{A}\cap\mathfrak{B}=\mathfrak{A}\mathfrak{B}+\mathfrak{B}\mathfrak{A}$.

For then $\mathfrak{A}\cap\mathfrak{B}=(\mathfrak{A}+\mathfrak{B})(\mathfrak{A}\cap\mathfrak{B})=\mathfrak{A}(\mathfrak{A}\cap\mathfrak{B})+\mathfrak{B}(\mathfrak{A}\cap\mathfrak{B})\subseteq\mathfrak{A}\mathfrak{B}+\mathfrak{B}\mathfrak{A}\subseteq\mathfrak{A}\cap\mathfrak{B}$.

7.13. $[\mathfrak{A}_1,\mathfrak{A}_2]\,\mathfrak{B}\subseteq[\mathfrak{A}_1\mathfrak{B},\mathfrak{A}_2\mathfrak{B}]$.

As regards associativity of ideal multiplication, we can only say:

7.14. $(\mathfrak{A}\mathfrak{B})\mathfrak{C}\subseteq\mathfrak{A}(\mathfrak{B}\mathfrak{C})$ *for any three ideals of* Γ.

Proof. The product $(\mathfrak{A}\mathfrak{B})\mathfrak{C}$ consists, by 7.07, of all finite sums whose terms are of the form $(\sum\alpha_i\beta_i)\gamma=\sum\alpha_i\beta_i\gamma$; hence it consists of all sums of the form $\sum\alpha\beta\gamma$. The product $\mathfrak{A}(\mathfrak{B}\mathfrak{C})$ consists of all sums of the form $\sum_i\alpha_i\left(\sum_j\beta_{i,j}\gamma_{i,j}\right)$. Clearly the sums $\sum\alpha\beta\gamma$ which are the elements of $(\mathfrak{A}\mathfrak{B})\mathfrak{C}$, occur amongst these.

We shall see (9.12) that in a near-ring Γ_n of finite rank the inclusion may be proper. Thus ideal multiplication need not be associative, in fact not even power associative. The restriction to a finite number of generators will prove the reason for the non-associativity; this is born out by the fact that in Γ_ω ideal multiplication is associative:

7.15. Theorem. *For any three ideals* $\mathfrak{A},\mathfrak{B}$, *and* $\mathfrak{C}$ *of* Γ_ω

$$(\mathfrak{A}\mathfrak{B})\mathfrak{C}=\mathfrak{A}(\mathfrak{B}\mathfrak{C}).$$

This follows at once from

7.16. Theorem. *The product* $\mathfrak{A}\mathfrak{B}$ *of two ideals of* Γ_ω *consists of all products* $\alpha\beta$, $\alpha\in\mathfrak{A}$, $\beta\in\mathfrak{B}$.

Again, by means of 7.07, this follows by repeated application of:

7.17. *If* $\mathfrak{A},\mathfrak{B}$ *are ideals of* Γ_ω, *and* α_1, α_2 *are elements of* $\mathfrak{A}$, β_1,β_2 *are elements of* $\mathfrak{B}$, *then there exist elements* $\alpha\in\mathfrak{A}$, $\beta\in\mathfrak{B}$, *such that* $\alpha_1\beta_1+\alpha_2\beta_2=\alpha\beta$.

Proof. We introduce for every integer n some special endomorphisms of Γ_ω. Let σ_n be the endomorphism given by

$$e_i\sigma_n=e_{i+n}\quad\text{for}\quad i=1,\dots,n,$$
$$e_i\sigma_n=0\quad\text{for}\quad i>n.$$

Let τ_n be the endomorphism given by

$$e_i\tau_n=0\quad\text{for}\quad i=1,\dots,n,$$
$$e_i\tau_n=e_{i-n}\quad\text{for}\quad i=n-1,\dots,2n,$$
$$e_i\tau_n=0\quad\text{for}\quad i>2n.$$

Right multiplication by σ_n replaces the generators $e_1,\dots,e_n$ in an element $\alpha\in\Gamma_n\subset\Gamma_\omega$ by $e_{n+1},\dots,e_{2n}$. Left multiplication by τ_n of an element $\alpha\in\Gamma_n\subset\Gamma_\omega$ has the effect of moving the first n components of α "down" to form the $(n+1)^{\text{st}},\dots,(2n)^{\text{th}}$ components, and replacing the first n components by zeros.

Now let n be the least integer such that the given elements α_1, α_2 and β_1, β_2 all lie in the sub-near-ring Γ_n of Γ_ω. Then

$$7.18. \qquad \alpha_1\beta_1 + \alpha_2\beta_2 = (\alpha_1 + \alpha_2\sigma_n)(\beta_1 + \tau_n\beta_2),$$

where

$$\alpha_1 + \alpha_2\sigma_n = \alpha \in \mathfrak{A}, \quad \text{and} \quad \beta_1 + \tau_n\beta_2 = \beta \in \mathfrak{B}.$$

This is easily checked by means of the rules 4.04 for calculating the components of the sum and product of endomorphisms, observing that the four elements $\alpha_1, \alpha_2, \beta_1, \beta_2$ all have only n non-trivial components, which depend on $e_1, \ldots, e_n$ only.

The proof clearly depends on the possibility to use an unlimited number of generators.

A number of questions naturally arise out of these last theorems which seem to make ideal multiplication in Γ_ω quite tractable; there are in particular the problems arising out of the notion of ideal bases. One might hope that at least in Γ_ω the product of two finitely based ideals is itself finitely based. However, what little information is available, consists of isolated examples treated by ad hoc methods. A brief summary of these is given in the Appendix.

It remains to interpret ideal multiplication in the free group associated with Γ.

Notation. The fully invariant subgroup corresponding to the product $\mathfrak{A}\mathfrak{B}$ of two ideals of Γ is denoted by $A \circ B$.

The rule 4.04 for forming the product of two elements of the near-ring shows that $A \circ B$ is the fully invariant subgroup obtained by substituting the elements of B for the generators in the words representing the elements of A and taking the additive closure. Moreover:

7.19. Theorem. *The fully invariant subgroup $A \circ B$ of the free group associated with Γ corresponding to the ideal product $\mathfrak{A}\mathfrak{B}$ in Γ is a fully invariant subgroup of B.*

Proof. As $\mathfrak{A}\mathfrak{B} \subseteqq \mathfrak{B}$, certainly $A \circ B \subseteqq B$; now $A \circ B$ is generated by all elements of the form $a(b_1, \ldots, b_n)$ [6], where $a(e_1, \ldots, e_n) \in A$, and $b_1, \ldots, b_n \in B$. But the set of all such elements is closed under endomorphisms of B; for any such endomorphism maps the set $b_1, \ldots, b_n$ onto some set of elements $b_1', \ldots, b_n' \in B$, and as it is a homomorphic mapping, it maps the word $a(b_1, \ldots, b_n)$ onto the word $a(b_1', \ldots, b_n')$, that is on an element of the generating set of $A \circ B$.

The converse of this last theorem is not true in general, as will be seen in the last paragraph; again the restriction to a finite number of generators in the near-rings Γ_n is the reason.

8. The lattice of ideals and ideal multiplication in Φ_ω

From now on we consider the near-rings associated with the absolutely free groups F_n and F only. All we really need is that the free groups we use

have the SCHREIER property that every subgroup is itself a free group. However, this further restriction is not likely to make much difference: we are not concerned with the case of abelian varieties, that is, the theory of ordinary matrix rings. But no non-abelian free groups with the SCHREIER property other than the absolutely free groups are known at present. On the other hand, much of the general picture given in the next two paragraphs remains true also for the near-rings associated with arbitrary varieties; the differences are caused, roughly speaking, by the homomorphism of the near-rings Φ onto the near-rings Γ obliterating distinctions that exist in the near-rings associated with the absolutely free groups.

The semigroup of ideals in Φ_ω is our first concern; the results will be used in the next paragraph to clarify the reason for the non-associativity of ideal multiplication in Φ_n by "reduction to Φ_n" in the sense of § 2.

We show first that in Φ_ω Theorem 7.19 possesses a converse:

8.01. *If B and C are fully invariant subgroups of F, and C is a fully invariant subgroup of B, if further $\mathfrak{B}$ and $\mathfrak{C}$ are the ideals of Φ_ω corresponding to B and C, then there exists an ideal $\mathfrak{A}$ such that $\mathfrak{C} = \mathfrak{A}\mathfrak{B}$.*

Proof. Since C is fully invariant in B, it is a word subgroup of the free group B. This means: if $b_1, b_2, \ldots$ form a set of free generators of B, there exists a set of words $a_i(b_j)$ in these free generators such that C is the fully invariant closure in B of this set of words. Consider the corresponding words $a_i(e_j)$ of F; let A be their fully invariant closure in F, and $\mathfrak{A}$ the ideal of Φ_ω corresponding to A. Then one confirms easily that $\mathfrak{C} = \mathfrak{A}\mathfrak{B}$. One sees why this need not be true in Φ_n: the words defining C as word subgroup of B may involve more than n generators.

The proof suggets that $\mathfrak{A}$ is uniquely determined by $\mathfrak{B}$ and $\mathfrak{C}$. To confirm this, we return to the free generators $b_1, b_2, \ldots$ of the fully invariant subgroup B. The set b_i is necessarily infinite if $\mathfrak{B} \neq \{0\}$; we denote by β_m the element of Φ_ω given by $e_i\beta_m = b_i$ for $i = 1, \ldots, m$, $e_i\beta_m = 0$ for $i > m$. If β is any element of $\mathfrak{B}$, its finitely many non-trivial components are representable as words in finitely many b_i, and this representation is unique. Hence:

8.02. *Every $\beta \in \mathfrak{B}$ is representable in the form $\beta = \xi\beta_m$ for some m and an element $\xi \in \Phi_\omega$ whose components belong to F_m.*

Moreover, since the b_i are free generators of B:

8.03. *If ξ and η are elements of Φ_ω whose components belong to F_m and F_n respectively, then*

$$\xi\beta_m = \eta\beta_n \text{ implies } \xi = \eta.$$

Now we obtain as a first step:

8.04. *If $\mathfrak{A}_1 \geqq \mathfrak{A}_2$, and $\mathfrak{A}_1\mathfrak{B} = \mathfrak{A}_2\mathfrak{B}$, where $\mathfrak{B} \neq \{0\}$, then $\mathfrak{A}_1 = \mathfrak{A}_2$.*

Proof. Let α_1 be an arbitrary element of $\mathfrak{A}_1$. There exists an m such that the components of α_1 all belong to F_m. We form the product $\alpha_1\beta_m$. It belongs to $\mathfrak{A}_1\mathfrak{B} = \mathfrak{A}_2\mathfrak{B}$; but as element of the latter, it is of the form (by 7.16):

$$\alpha_1\beta_m = \alpha_2\beta \quad \text{with} \quad \alpha_2 \in \mathfrak{A}_2,\ \beta \in \mathfrak{B}.$$

Now we use 8.02 to write β in the form $\beta = \eta \beta_n$ with an element η whose components lie in F_n. But then also $\alpha_2 \eta$ has components in F_n, and

$$\alpha_1 \beta_m = \alpha_2 \beta = \alpha_2 \eta \beta_n,$$

hence by 8.03, $\alpha_1 = \alpha_2 \eta \in \mathfrak{A}_2$, and so $\mathfrak{A}_1 \subseteq \mathfrak{A}_2$, giving $\mathfrak{A}_1 = \mathfrak{A}_2$.

To complete the argument, we introduce the right quotient of any two ideals $\mathfrak{B}$ and $\mathfrak{C}$.

8.05. Definition[7]). *The right quotient* $\mathfrak{C} r \mathfrak{B}$ (*read:* $\mathfrak{C}$ *divided on the right by* $\mathfrak{B}$) *of any two ideals* $\mathfrak{C}$ *and* $\mathfrak{B}$ *is the set of all elements* $\varrho \in \Phi_\omega$ *such that* $\varrho \beta \in \mathfrak{C}$ *for all* $\beta \in \mathfrak{B}$.

A straightforward calculation shows:

8.06. $\mathfrak{C} r \mathfrak{B}$ *is an ideal containing* $\mathfrak{C}$.

If $\mathfrak{B}_1 \subseteq \mathfrak{B}_2$, *then* $\mathfrak{C} r \mathfrak{B}_1 \supseteq \mathfrak{C} r \mathfrak{B}_2$; $\quad \mathfrak{C} r \Phi_\omega = \mathfrak{C}$.

If $\mathfrak{C}_1 \subseteq \mathfrak{C}_2$, *then* $\mathfrak{C}_1 r \mathfrak{B} \subseteq \mathfrak{C}_2 r \mathfrak{B}$.

8.07. Theorem. *If* $\mathfrak{R} = \mathfrak{C} r \mathfrak{B}$, *then* $\mathfrak{R}\mathfrak{B} \subseteq \mathfrak{B} \cap \mathfrak{C}$, *and* $R \circ B$ *is the greatest fully invariant subgroup of* B *contained in* $B \cap C$.

Proof. $\mathfrak{R}$ is an ideal by 8.06, and $\mathfrak{R}\mathfrak{B} \subseteq \mathfrak{B}$ since a product is contained in each factor. Also $\mathfrak{R}\mathfrak{B} \subseteq \mathfrak{C}$ by definition of $\mathfrak{R}$, hence $\mathfrak{R}\mathfrak{B} \subseteq \mathfrak{B} \cap \mathfrak{C}$. By 7.19, $R \circ B$ is fully invariant in B. Let D be any fully invariant subgroup of B contained in $B \cap C$, and let $\mathfrak{D}$ be the corresponding ideal. By 8.01 there exists an ideal $\mathfrak{A}$ such that $\mathfrak{D} = \mathfrak{A}\mathfrak{B}$. Since $\mathfrak{D} \subseteq \mathfrak{C}$, $\mathfrak{A} \subseteq \mathfrak{R}$ by definition of the right quotient. Hence $\mathfrak{D} = \mathfrak{A}\mathfrak{B} \subseteq \mathfrak{R}\mathfrak{B}$, and so $D = A \circ B \subseteq R \circ B$, that is, $R \circ B$ is maximal.

If now, in particular, C is a fully invariant subgroup of B, then $(\mathfrak{C} r \mathfrak{B})\mathfrak{B} = \mathfrak{C}$; and if also $\mathfrak{A}\mathfrak{B} = \mathfrak{C}$, then $\mathfrak{A} \subseteq \mathfrak{C} r \mathfrak{B}$, hence by 8.04: $\mathfrak{A} = \mathfrak{C} r \mathfrak{B}$. Thus we have proved:

8.08. Theorem. *In* Φ_ω, $\mathfrak{C} = \mathfrak{A}\mathfrak{B}$ *with* $\mathfrak{B} \neq \{0\}$ *if, and only if,* C *is a fully invariant subgroup of* B. *Here* $\mathfrak{A}$ *is uniquely determined by* $\mathfrak{B}$ *and* $\mathfrak{C}$, *namely* $\mathfrak{A} = \mathfrak{C} r \mathfrak{B}$.

We obtain as corollary:

8.09. Theorem. *In* Φ_ω, $\mathfrak{A}_1\mathfrak{B} = \mathfrak{A}_2\mathfrak{B}$, *where* $\mathfrak{B} \neq \{0\}$, *implies* $\mathfrak{A}_1 = \mathfrak{A}_2$.

The fact that the ideals in Φ_ω form a semigroup with right cancellation suggests carrying the investigation of right multiplication by a fixed ideal a little further. We know already (7.11) that

$$(\mathfrak{A}_1 + \mathfrak{A}_2)\mathfrak{B} = \mathfrak{A}_1\mathfrak{B} + \mathfrak{A}_2\mathfrak{B}.$$

We now prove:

8.10. *In* Φ_ω, $\qquad (\mathfrak{A}_1 \cap \mathfrak{A}_2)\mathfrak{B} = \mathfrak{A}_1\mathfrak{B} \cap \mathfrak{A}_2\mathfrak{B}$.

[7]) The definition should reasonably be extended to any left ideal $\mathfrak{C}$ and its right quotient by the ideal $\mathfrak{B}$; but we make no use of this. The subsequent statements 8.06 and 8.07 remain true, except that in this more general case $\mathfrak{C} r \mathfrak{B}$ need not contain $\mathfrak{C}$. A specific example: $\mathfrak{C} r \Phi_\omega$ corresponds to the greatest fully invariant subgroup of F contained in the normal subgroup corresponding to the left ideal $\mathfrak{C}$ by 6.05.

Proof. Since the right-hand side contains the left-hand side (7.10), we need only prove $\mathfrak{A}_1\mathfrak{B} \cap \mathfrak{A}_2\mathfrak{B} \subseteq (\mathfrak{A}_1 \cap \mathfrak{A}_2)\mathfrak{B}$. Let γ be any element of $\mathfrak{A}_1\mathfrak{B} \cap \mathfrak{A}_2\mathfrak{B}$. By 7.16, γ possesses representations

$$\gamma = \alpha_1\beta' = \alpha_2\beta'', \quad \text{where} \quad \alpha_1 \in \mathfrak{A}_1,\ \alpha_2 \in \mathfrak{A}_2,\ \beta', \beta'' \in \mathfrak{B}.$$

We write the elements β' and β'' of $\mathfrak{B}$ in terms of free generators of B as in 8.02: $\beta' = \xi\beta_m$, $\beta'' = \eta\beta_n$, where ξ and η have components in F_m and F_n respectively. Then $\gamma = \alpha_1\xi\beta_m = \alpha_2\eta\beta_n$, where $\alpha_1\xi$ and $\alpha_2\eta$ also have components in F_m and F_n respectively. Hence, by 8.03, $\alpha_1\xi = \alpha_2\eta$. But $\alpha_1\xi \in \mathfrak{A}_1$, $\alpha_2\eta \in \mathfrak{A}_2$, hence $\alpha_1\xi = \alpha_2\eta = \alpha \in \mathfrak{A}_1 \cap \mathfrak{A}_2$, and so we have $\gamma = \alpha\beta_m \in (\mathfrak{A}_1 \cap \mathfrak{A}_2)\mathfrak{B}$. Therefore $\mathfrak{A}_1\mathfrak{B} \cap \mathfrak{A}_2\mathfrak{B} \subseteq (\mathfrak{A}_1 \cap \mathfrak{A}_2)\mathfrak{B}$, and 8.10 follows.

We note that a similar argument will prove also $[\mathfrak{A}_1\mathfrak{B}, \mathfrak{A}_2\mathfrak{B}] = [\mathfrak{A}_1, \mathfrak{A}_2]\mathfrak{B}$; but we shall not make use of this. We sum up the preceding results in:

8.11. Theorem. *In the near-ring Φ_ω right multiplication by a fixed ideal $\mathfrak{B} \neq \{0\}$ constitutes an iso-endomorphism of the lattice of all ideals. This iso-endomorphism corresponds in F to the mapping of the lattice of all fully invariant subgroups, that is word subgroups, of F onto the sub-lattice of all word subgroups of the fully invariant subgroup B of F.*

An ideal of Φ_ω is a non-trivial product of two ideals if, and only if, the corresponding fully invariant subgroup of F is fully invariant in a proper fully invariant supergroup. We say:

8.12. Definition. *An ideal of Φ_ω is called indecomposable, if it is not the product of two proper ideals of Φ_ω.*

Examples of indecomposable ideals are all maximal ideals, but also, for instance, the derived ideal.

8.13. *Every ideal of Φ_ω is representable as a product of indecomposable ideals.*

Proof. If the ideal $\mathfrak{A}$ is not itself indecomposable, then $\mathfrak{A} = \mathfrak{B}_1\mathfrak{B}_2$. If at least one of the factors is not indecomposable, one factorizes further. That this process must break off after a finite number of steps, follows from LEVI's theorem on chains of characteristic subgroups of a free group (cf. [7]): If $A = A_t \subset A_{t-1} \subset \cdots \subset A_0 = F$ is a chain of characteristic subgroups of F, and each term is characteristic in its successor, then the length t of the chain is at most equal to $l(A) - 2$, where $l(A)$ is the length (qua reduced word in F) of a shortest element of A. For this length $l(A_i)$ increases by at least 1 every time one descends from A_{i-1} to A_i in the chain.

A representation $\mathfrak{A} = \mathfrak{B}_1\mathfrak{B}_2 \ldots \mathfrak{B}_t$ of $\mathfrak{A}$ as a product of t factors leads to a chain of fully invariant subgroups of F, namely

$$A \subset B_2 \circ B_3 \circ \cdots \circ B_i \subset B_3 \circ \cdots \circ B_t \subset \cdots \subset B_{t-1} \circ B_t \subset B_t \subset F.$$

Every term is fully invariant, hence a fortiori characteristic, in its successor. It follows that t, that is the number of factors of $\mathfrak{A}$, is bounded by $\mathfrak{A}$ alone.

Nothing is known on the uniqueness, or otherwise, of a representation as a product of indecomposable factors. In fact it is not even sure that the left cancellation law holds in the semigroup of ideals. In the remainder of this paragraph we collect some partial results in this direction.

8.14. *If* $\mathfrak{A} \neq \Phi_\omega$, $\{0\}$, *then* $\mathfrak{A}^{n+1} < \mathfrak{A}^n$ *for all* n, *and* $\bigcap_n \mathfrak{A}^n = \{0\}$.

Proof. Certainly $\mathfrak{A}^{n+1} \leqq \mathfrak{A}^n$. From $\mathfrak{A}^{n+1} = \mathfrak{A}\mathfrak{A}^n = \mathfrak{A}^n = \Phi_\omega \mathfrak{A}^n$ follows $\mathfrak{A} = \Phi_\omega$ by 8.09. Hence $\mathfrak{A}^{n+1} < \mathfrak{A}^n$. The fully invariant subgroup belonging to $\mathfrak{A}^{n+1}$ is fully invariant in that belonging to $\mathfrak{A}^n$. Hence, again by LEVI's theorem quoted above, the descending chain of these fully invariant subgroups contracts to $\{0\}$.

We deduce:

8.15. *If* $\mathfrak{B} \neq \{0\}$, *and* $\mathfrak{A}\mathfrak{B}\mathfrak{C} = \mathfrak{B}$, *then* $\mathfrak{A} = \mathfrak{C} = \Phi_\omega$.

Proof. From $\mathfrak{A}\mathfrak{B}\mathfrak{C} = \mathfrak{B}$ one has inductively: $\mathfrak{A}^n\mathfrak{B}\mathfrak{C}^n = \mathfrak{B}$ for all n. Hence $\mathfrak{B} \leqq \mathfrak{A}^n \cap \mathfrak{C}^n$ for all n; thus, by 8.14, $\mathfrak{B} = \{0\}$ unless $\mathfrak{A} = \mathfrak{C} = \Phi_\omega$.

Now 8.08, 8.09, and 8.15 give immediately:

8.16. *If* $\mathfrak{B}_1 \geqq \mathfrak{B}_2$, *and* B_2 *is fully invariant in* B_1, *then* $\mathfrak{A}\mathfrak{B}_1 = \mathfrak{A}\mathfrak{B}_2$ *with* $\mathfrak{A} \neq \{0\}$ *implies* $\mathfrak{B}_1 = \mathfrak{B}_2$.

A little more can be said, if $\mathfrak{B}_1$ and $\mathfrak{B}_2$ generate the whole of Φ_ω. In that case, $\mathfrak{A}\mathfrak{B}_1 = \mathfrak{A}\mathfrak{B}_2$ can hold—if at all—only for an ideal $\mathfrak{A}$ which is contained in $\mathfrak{B}_1 \cap \mathfrak{B}_2$, and properly contained in the derived ideal.

8.17. *If* $\mathfrak{B}_1, \mathfrak{B}_2 \neq \Phi_\omega$, $\mathfrak{B}_1 + \mathfrak{B}_2 = \Phi_\omega$, *and* $\mathfrak{A}\mathfrak{B}_1 = \mathfrak{A}\mathfrak{B}_2$, *then* $\mathfrak{A} \leqq \mathfrak{B}_1 \cap \mathfrak{B}_2$.

Proof. By 7.12, we have $\mathfrak{A} = \mathfrak{A}(\mathfrak{B}_1 + \mathfrak{B}_2) \leqq \mathfrak{A}\mathfrak{B}_1 + \mathfrak{A}\mathfrak{B}_2 + (\mathfrak{A} \cap \mathfrak{B}_1 \cap \mathfrak{B}_2)$. Since $\mathfrak{A}\mathfrak{B}_1 \leqq \mathfrak{A} \cap \mathfrak{B}_1$, $\mathfrak{A}\mathfrak{B}_2 \leqq \mathfrak{A} \cap \mathfrak{B}_2$, it follows that $\mathfrak{A}\mathfrak{B}_1 = \mathfrak{A}\mathfrak{B}_2 \leqq \mathfrak{A} \cap \mathfrak{B}_1 \cap \mathfrak{B}_2$, hence $\mathfrak{A} \leqq \mathfrak{A} \cap \mathfrak{B}_1 \cap \mathfrak{B}_2$, which means $\mathfrak{A} \leqq \mathfrak{B}_1 \cap \mathfrak{B}_2$.

For the proof of the second statement above, we use the following

Notation: The derived ideal, that is the ideal corresponding to the derived group of F, is called $\mathfrak{D}$. The ideal corresponding to the fully invariant closure of the single word $m e_1$, that is the fully invariant subgroup generated by all m^{th} multiples of elements of F, is called $\mathfrak{E}_m$.

8.18. *If* $\mathfrak{B}_1, \mathfrak{B}_2 \neq \Phi_\omega$, $\mathfrak{B}_1 + \mathfrak{B}_2 = \Phi_\omega$, *and* $\mathfrak{A}\mathfrak{B}_1 = \mathfrak{A}\mathfrak{B}_2$, *then* $\mathfrak{A} < \mathfrak{D}$.

Proof. If $\mathfrak{B}_1$ and $\mathfrak{B}_2$ together generate the whole of Φ_ω, $\mathfrak{B}_1$ must contain $\mathfrak{E}_m$, and $\mathfrak{B}_2$ must contain $\mathfrak{E}_n$, for some integers m and n which are relative prime. We assume m and n to be as small as possible with these properties.

1. Assume $\mathfrak{A}$ contains elements outside the ideal $\mathfrak{D}$. Then $\mathfrak{A}$ contains the ideal $\mathfrak{E}_p$ for some positive integer p; again we take p as small as possible. Then $\mathfrak{A}\mathfrak{B}_1$ contains $\mathfrak{E}_{mp}$, and $\mathfrak{A}\mathfrak{B}_2$ contains $\mathfrak{E}_{np}$, and in each case, mp and np respectively are the least integers with this property. As $p \neq 0$, $mp \neq np$, hence $\mathfrak{A}\mathfrak{B}_1 \neq \mathfrak{A}\mathfrak{B}_2$ contrary to our assumption. Incidentally, in this case only $m \neq n$ has been used.

2. It remains to consider the case $\mathfrak{A} = \mathfrak{D}$. Here it suffices to show that $\mathfrak{D}\mathfrak{B}_1 + \mathfrak{D}\mathfrak{B}_2 = \mathfrak{D}$. For then $\mathfrak{D}\mathfrak{B}_1 = \mathfrak{D}\mathfrak{B}_2$ implies $\mathfrak{D}\mathfrak{B}_1 = \mathfrak{D}\mathfrak{B}_2 = \mathfrak{D}$, hence $\mathfrak{B}_1 = \mathfrak{B}_2 = \Phi_\omega$ by 8.15, contrary to our assumption.

Now $\mathfrak{E}_m \subseteqq \mathfrak{B}_1$, and $\mathfrak{E}_n \subseteqq \mathfrak{B}_2$, hence $\mathfrak{D}\mathfrak{E}_m + \mathfrak{D}\mathfrak{E}_n \subseteqq \mathfrak{D}\mathfrak{B}_1 + \mathfrak{D}\mathfrak{B}_2$. Therefore it is sufficient to prove:

8.19. *If the integers m and n are relative prime, then* $\mathfrak{D}\mathfrak{E}_m + \mathfrak{D}\mathfrak{E}_n = \mathfrak{D}$.

Proof[8]). In terms of the corresponding fully invariant subgroups of F, 8.19 states that the derived groups of the fully invariant subgroups of F which are generated by all m^{th} multiples, and all n^{th} multiples of elements of F respectively, together generate the whole derived group of F. This will be proved by showing: Any group H, in which all m^{th} multiples permute with each other, and all n^{th} multiples permute with each other, where m and n are relative prime, is an abelian group.

Under the assumptions, the mth multiples of elements of H generate an abelian normal subgroup M of H. The factor group H/M is of exponent m; therefore every element of H/M is an nth multiple, as m is prime to n. But n^{th} multiples commute in H, and a fortiori also in H/M. Thus H/M is abelian, and M contains the derived group H_1 of H. By symmetry, H_1 is also contained in the normal subgroup N of H generated by all n^{th} multiples of elements of H, and so $H_1 \subseteqq M \cap N$. But $M \cap N$ consists of elements permuting with both, all m^{th} and all n^{th} multiples, and these are in the centre of H, as every element h of H can be written in the form $h = r(mh) + s(nh)$ for suitable integers r and s. Hence the derived group belongs to the centre. If now $c = [a, b]$ is any commutator of H, then $[pa, qb] = pqc$ for all integers p and q. But $0 = [ma, mb] = m^2c$, and $0 = [na, nb] = n^2c$. Since m and n are relative prime, $c = 0$ follows, and so H is abelian.

Theorem 8.11 permits us to extend the results 8.17 and 8.18 slightly, without much significance for our knowl'dge of the semigroup of ideals of Φ_ω.

The obvious question whether there exist representations of ideals, or of indecomposable ideals, as intersections of larger ideals also remains unanswered. The missing distributive law, but more than that, the absence of a descending chain condition, and the lack of knowledge on ascending chain conditions make it doubtful that this approach befits the problem.

9. The relation between ideal multiplication in Φ_n and Φ_ω

We use the same nomenclature as in § 2: If $\mathfrak{A}$ is an ideal of Φ_ω, or of Φ_n, the intersection of $\mathfrak{A}$ with the sub-near-ring Φ_m $(n > m)$ is an ideal of Φ_m, which we call the restriction of $\mathfrak{A}$ to Φ_m. Similarly, if $\mathfrak{A}$ is an ideal in Φ_m, the ideal generated by $\mathfrak{A}$ in Φ_n, or in Φ_ω, is the extension of $\mathfrak{A}$ to Φ_n, or to Φ_ω, respectively. These notions correspond, of course, exactly to those used in § 2:

9.01. *The fully invariant subgroup belonging to the restriction (extension) of an ideal is the restriction (extension) of the fully invariant subgroup belonging to that ideal.*

[8]) This proof is due to my husband.

Hence, the statements 2.02 to 2.06 which govern these processes of restriction and extension hold, mutatis mutandis, for ideals in the near-rings.

Notation. If $\mathfrak{A}$ is an ideal of Φ_ω, its restriction to Φ_n is denoted by $\mathfrak{A}_n$, $\mathfrak{A}_n = \mathfrak{A} \cap \Phi_n$. The extension of $\mathfrak{A}_n$ to Φ_ω is denoted by $\mathfrak{A}^{(n)}$.

We summarize the facts we are going to use here:

9.02. *The ideals $\mathfrak{A}^{(n)}$ of Φ_ω form an ascending chain,*

$$\mathfrak{A}^{(1)} \leqq \mathfrak{A}^{(2)} \leqq \cdots \leqq \mathfrak{A}^{(n)} \leqq \cdots \leqq \mathfrak{A}.$$

Here $\mathfrak{A}^{(n)} = \mathfrak{A}$ if, and only if, $\mathfrak{A}$ can be generated by elements of Φ_n. The ideals $\mathfrak{A}$ and $\mathfrak{A}^{(n)}$ have the same restriction $\mathfrak{A}_n$ to Φ_n.

To form the restriction $\mathfrak{A}_n$ of $\mathfrak{A}$ to Φ_n, we may use the relative identity element ε_n to pick out, as in 5.03, that part of $\mathfrak{A}$ that lies in Φ_n:

$$\mathfrak{A}_n = \mathfrak{A} \cap \Phi_n = \mathfrak{A} \cap \varepsilon_n \Phi_\omega \varepsilon_n = \varepsilon_n \mathfrak{A} \varepsilon_n.$$

In particular:

9.03. *The restriction $(\mathfrak{A}\mathfrak{B})_n$ of the product $\mathfrak{A}\mathfrak{B}$ to Φ_n consists of all elements of the form $(\varepsilon_n \alpha)(\beta \varepsilon_n)$ with $\alpha \in \mathfrak{A}$, $\beta \in \mathfrak{B}$.*

For, by 7.16, $\mathfrak{A}\mathfrak{B}$ consists of all products $\alpha\beta$, $\alpha \in \mathfrak{A}$, $\beta \in \mathfrak{B}$; hence $\mathfrak{A}\mathfrak{B} \cap \Phi_n = \varepsilon_n \mathfrak{A}\mathfrak{B}\varepsilon_n$ consists of all products $\varepsilon_n \alpha \beta \varepsilon_n$. The factor $\varepsilon_n \alpha$ is an element with at most n non-trivial components, which belong to A. The second factor $\beta\varepsilon_n$ is an element whose components belong to F_n, that is, to $B \cap F_n = B_n$: the number of non-trivial components may, of course, be greater than n. But as $B_n \subset B^{(n)}$, we always have:

9.04. *If $\beta \in \mathfrak{B}$, then $\beta\varepsilon_n \in \mathfrak{B}^{(n)}$.*

It follows:

9.05. *For any two ideals $\mathfrak{A}$ and $\mathfrak{B}$ of Φ_ω,*

$$\mathfrak{A}\mathfrak{B} \geqq \mathfrak{A} \cdot \mathfrak{B}^{(n)} \geqq (\mathfrak{A}\mathfrak{B})^{(n)},$$

and the restriction to Φ_n of all three terms is the same ideal of Φ_n.

Proof. As $\mathfrak{B}^{(n)} \leqq \mathfrak{B}$, clearly $\mathfrak{A}\mathfrak{B}^{(n)} \leqq \mathfrak{A}\mathfrak{B}$. Also, $(\mathfrak{A}\mathfrak{B})^{(n)}$ is the ideal generated in Φ_ω by the restriction of $\mathfrak{A}\mathfrak{B}$ to Φ_n; hence 9.03 shows that it is generated by all elements of the form $(\varepsilon_n \alpha)(\beta\varepsilon_n)$. But $\varepsilon_n \alpha \in \mathfrak{A}$, $\beta\varepsilon_n \in \mathfrak{B}^{(n)}$ by 9.04, hence every generator of $(\mathfrak{A}\mathfrak{B})^{(n)}$ belongs to $\mathfrak{A}\mathfrak{B}^{(n)}$, and so $(\mathfrak{A}\mathfrak{B})^{(n)} \leqq \mathfrak{A}\mathfrak{B}^{(n)}$. Further, by 9.02, $\mathfrak{A}\mathfrak{B}$ and $(\mathfrak{A}\mathfrak{B})^{(n)}$ have the same restriction to Φ_n, therefore the same is true for any ideal between these two.

9.06. *The restriction of $\mathfrak{A}^{(n)} \cdot \mathfrak{B}$ to Φ_n is the product $\mathfrak{A}_n\mathfrak{B}_n$ in Φ_n,*

$$\mathfrak{A}^{(n)}\mathfrak{B} \cap \Phi_n = \mathfrak{A}_n\mathfrak{B}_n.$$

Proof. As $\mathfrak{A}_n \subset \mathfrak{A}^{(n)}$ and $\mathfrak{B}_n \subset \mathfrak{B}$, $\mathfrak{A}_n\mathfrak{B}_n \leqq \mathfrak{A}^{(n)}\mathfrak{B}$; the product $\mathfrak{A}_n\mathfrak{B}_n$ trivially belongs to Φ_n, hence $\mathfrak{A}_n\mathfrak{B}_n \leqq \mathfrak{A}^{(n)} \cdot \mathfrak{B} \cap \Phi_n$. To prove the reverse inclusion, we note that $\mathfrak{A}^{(n)}$ is generated by $\mathfrak{A}_n$ in Φ_ω; hence, by 7.05, $\mathfrak{A}^{(n)}$ consists of all elements of the form $\sum \xi \alpha_n \eta$, where $\alpha_n \in \mathfrak{A}_n$ and $\xi, \eta \in \Phi_\omega$. Therefore, by 7.16, $\mathfrak{A}^{(n)}\mathfrak{B}$ consists of all elements of the form $\gamma = (\sum \xi \alpha_n \eta)\beta$.

Here $\alpha_n = \varepsilon_n \alpha_n \varepsilon_n$, as $\alpha_n \in \Phi_n$. If now the element γ belongs also to Φ_n, then $\gamma = \varepsilon_n \gamma \varepsilon_n$. Hence, $\gamma \in \mathfrak{A}^{(n)}\mathfrak{B} \cap \Phi_n$ implies:

$$\begin{aligned}\gamma &= \varepsilon_n(\sum \xi\, \varepsilon_n \alpha_n\, \varepsilon_n \eta)\, \beta\, \varepsilon_n \\ &= \sum \varepsilon_n \xi\, \varepsilon_n \cdot \alpha_n \cdot \varepsilon_n \eta\, \beta\, \varepsilon_n \quad \text{by 5.04,} \\ &= \sum \xi_n \alpha_n \beta_n, \quad \text{where} \quad \xi_n \in \Phi_n, \beta_n \in \varepsilon_n \mathfrak{B}\, \varepsilon_n = \mathfrak{B}_n.\end{aligned}$$

Hence $\gamma \in \mathfrak{A}_n \mathfrak{B}_n$, which was to be proved.

Now 9.05, applied to $\mathfrak{A}^{(n)}$ in place of $\mathfrak{A}$ gives:

9.07. *For any two ideals $\mathfrak{A}$ and $\mathfrak{B}$ of Φ_ω,*

$$\mathfrak{A}^{(n)}\mathfrak{B} \supseteq \mathfrak{A}^{(n)}\mathfrak{B}^{(n)} \supseteq (\mathfrak{A}_n \mathfrak{B}_n)^{(n)},$$

and the restriction of all three terms to Φ_n is the same ideal $\mathfrak{A}_n \mathfrak{B}_n$.

This shows that the restriction to Φ_n of a product, whose first factor can be generated by elements involving at most n generators, is the product of the restrictions of the factors, irrespective of the nature of the second factor. The converse is not as straightforward: The rank $N(B_n)$ of the free subgroup B_n of F_n makes a difference to the situation. If $N = N(B_n)$ is finite, then, by SCHREIER's formula, $N \geq n$; and $N = n$ only if $B_n = F_n$. Here we have excluded the trivial case $B_n = \{0\}$. Therefore, whether N is finite or infinite, we have $\mathfrak{A} \supseteq \mathfrak{A}^{(N)} \supseteq \mathfrak{A}^{(n)}$, hence also $\mathfrak{A}\mathfrak{B} \supseteq \mathfrak{A}^{(N)}\mathfrak{B} \supseteq \mathfrak{A}^{(n)}\mathfrak{B}$[9]), and the restrictions of these products to Φ_n are in the same relation to each other. We show first:

9.08. *If B_n is of finite rank N as free subgroup of F_n, then*

$$\mathfrak{A}\mathfrak{B} \cap \Phi_n = \mathfrak{A}^{(N)}\mathfrak{B} \cap \Phi_n.$$

Proof. The right-hand side is contained in the left-hand side, hence we need only prove that $\mathfrak{A}\mathfrak{B} \cap \Phi_n \subseteq \mathfrak{A}^{(N)}\mathfrak{B} \cap \Phi_n$.

Let $b_1, \ldots, b_N$ $(N \geq n)$ be free generators of B_n, and let β_N denote the element of $\mathfrak{B}$ whose first N components are $b_1, \ldots, b_N$, and whose further components are zero. Then every element of $\mathfrak{B}$ whose components belong to B_n is representable in the form $\xi\, \beta_N$, where the components of ξ belong to F_N, that is $\xi = \xi\, \varepsilon_N$. If now $\gamma \in \mathfrak{A}\mathfrak{B} \cap \Phi_n$, then $\gamma = (\varepsilon_n \alpha)\, (\beta\, \varepsilon_n)$ for some $\alpha \in \mathfrak{A}$ and $\beta \in \mathfrak{B}$. But $\beta\, \varepsilon_n$ has components in B_n, hence $\gamma = (\varepsilon_n \alpha)\, (\xi\, \beta_N)$, where ξ has components in F_N; thus $\gamma = (\varepsilon_n \alpha)\, (\xi\, \varepsilon_N \beta_N) = (\varepsilon_n \alpha \xi\, \varepsilon_N)\, \beta_N$. Since $\varepsilon_n \alpha \xi \in \mathfrak{A}$, 9.04 shows that $\varepsilon_n \alpha\, \xi\, \varepsilon_N \in \mathfrak{A}^{(N)}$; hence $\gamma \in \mathfrak{A}^{(N)}\mathfrak{B}$. But $\gamma \in \Phi_n$, and so $\gamma \in \mathfrak{A}^{(N)}\mathfrak{B} \cap \Phi_n$, which proves 9.08.

Now we can complete the picture by showing:

9.09. Theorem. *If $\mathfrak{A}$ and $\mathfrak{B} \neq \{0\}$ are ideals of Φ_ω, where $B \cap F_n = B_n$ is of finite or infinite rank N as free subgroup of F_n, then the restriction to Φ_n of the product $\mathfrak{A}\mathfrak{B}$ is the product of the restrictions of $\mathfrak{A}$ and $\mathfrak{B}$, $\mathfrak{A}\mathfrak{B} \cap \Phi_n = \mathfrak{A}_n \mathfrak{B}_n$, if and only if, $\mathfrak{A}^{(N)} = \mathfrak{A}^{(n)}$.*

[9]) Here, and repeatedly further on, we use the natural convention: if N is infinite, then $\mathfrak{A}^{(N)} = \mathfrak{A}$.

Proof. If $\mathfrak{A}^{(N)} = \mathfrak{A}^{(n)}$, then 9.07 and 9.08 show that $\mathfrak{A}\mathfrak{B} \cap \Phi_n = \mathfrak{A}_n \mathfrak{B}_n$.

Conversely, $\mathfrak{A}\mathfrak{B} \cap \Phi_n = \mathfrak{A}_n \mathfrak{B}_n$ implies by 9.05, 9.07 and 9.08 that $\mathfrak{A}^{(N)}\mathfrak{B} \cap \Phi_n = \mathfrak{A}^{(n)}\mathfrak{B} \cap \Phi_n$. Since trivially $\mathfrak{A}^{(n)} \leqq \mathfrak{A}^{(N)}$, all we have to show is:

$$\mathfrak{A}^{(N)}\mathfrak{B} \cap \Phi_n = \mathfrak{A}^{(n)}\mathfrak{B} \cap \Phi_n \quad \text{implies} \quad \mathfrak{A}^{(N)} \leqq \mathfrak{A}^{(n)}.$$

1. Assume N finite. As $\mathfrak{A}^{(N)}$ is generated by elements α whose components lie in F_N, it is sufficient to show that every such element α of $\mathfrak{A}^{(N)}$ belongs to $\mathfrak{A}^{(n)}$. We consider, for any such α, the product $\alpha\beta_N \in \mathfrak{A}^{(N)}\mathfrak{B}$, where β_N is the element of $\mathfrak{B}$ whose non-trivial components are the N free generators of B_n (as in the proof of 9.08). Then

$$\varepsilon_n(\alpha\beta_N)\,\varepsilon_n = (\varepsilon_n\alpha)\,\beta_N \in \mathfrak{A}^{(N)}\mathfrak{B} \cap \Phi_n.$$

As $\mathfrak{A}^{(N)}\mathfrak{B} \cap \Phi_n = \mathfrak{A}^{(n)}\mathfrak{B} \cap \Phi_n$, there exist, by 9.03, elements $\alpha_1 \in \mathfrak{A}^{(n)}$ and $\beta_1 \in \mathfrak{B}$, such that $(\varepsilon_n\alpha)\beta_N = (\varepsilon_n\alpha_1)(\beta_1\varepsilon_n)$.

But $\beta_1\varepsilon_n$ has components in B_n, hence it is of the form $\beta_1\varepsilon_n = \xi\beta_N$, where the components of ξ lie in F_N; thus

$$(\varepsilon_n\alpha)\,\beta_N = (\varepsilon_n\alpha_1)\,(\xi\beta_N) = (\varepsilon_n\alpha_1\xi)\,\beta_N,$$

and this implies, by 8.03: $\varepsilon_n\alpha = \varepsilon_n\alpha_1\xi$.

Here $\varepsilon_n\alpha_1\xi$ lies in $\mathfrak{A}^{(n)}$, as α_1 belongs to $\mathfrak{A}^{(n)}$; hence the first n components of α, and therefore of every element of $\mathfrak{A}^{(N)}$ with components in F_N, belong to $A^{(n)}$. Since the components of the elements of an ideal vary independently over the corresponding fully invariant subgroup (cf. 6.05), every component of α belongs to $A^{(n)}$, and so $\alpha \in \mathfrak{A}^{(n)}$ as required.

2. Assume N infinite. We have to show:

$$\mathfrak{A}\mathfrak{B} \cap \Phi_n = \mathfrak{A}^{(n)}\mathfrak{B} \cap \Phi_n \quad \text{implies} \quad \mathfrak{A} \leqq \mathfrak{A}^{(n)}.$$

This time, let $b_1, b_2, \ldots$ be a set of free generators of B_n, and for every $m = 1, 2, \ldots$ denote by β_m the element of $\mathfrak{B}$ whose first m components are $b_1, \ldots, b_m$, and whose further components are zero. Consider any element $\alpha \in \mathfrak{A}$; the finitely many non-trivial components of α involve only a finite number of generators, $e_1, \ldots, e_k$ say. We consider the element $\alpha\beta_k$ of $\mathfrak{A}\mathfrak{B}$. Then $(\varepsilon_n\alpha)(\beta_k\varepsilon_n) = (\varepsilon_n\alpha)\beta_k$ is an element of $\mathfrak{A}\mathfrak{B} \cap \Phi_n$. Since $\mathfrak{A}\mathfrak{B} \cap \Phi_n = \mathfrak{A}^{(n)}\mathfrak{B} \cap \Phi_n$, there exist elements $\alpha_1 \in \mathfrak{A}^{(n)}$, $\beta_1 \in \mathfrak{B}$, such that $(\varepsilon_n\alpha)\beta_k = (\varepsilon_n\alpha_1)(\beta_1\varepsilon_n)$. Again, $\beta_1\varepsilon_n$ has components in B_n, hence is of the form $\beta_1\varepsilon_n = \xi\beta_m$ for some m, where ξ has components in F_m. Therefore $(\varepsilon_n\alpha)\beta_k = (\varepsilon_n\alpha_1)(\xi\beta_m) = (\varepsilon_n\alpha_1\xi)\beta_m$, which implies by 8.03, $\varepsilon_n\alpha = \varepsilon_n\alpha_1\xi \in \mathfrak{A}^{(n)}$. As before, $\mathfrak{A} \leqq \mathfrak{A}^{(n)}$ follows.

We are now also in a position to answer the question raised by Theorem 7.19: In every near-ring Φ_n or Φ_ω, the fully invariant subgroup $A \circ B$ is fully invariant in B. In Φ_ω, conversely, every fully invariant subgroup of B is of the form $A \circ B$ (Theorem 8.08). But in the near-rings Φ_n this is not true. Instead we have:

9.10. Theorem. *If B_n and C_n are fully invariant subgroups of F_n, and $C_n \leqq B_n$, then C_n is fully invariant in B_n if, and only if, the corresponding ideal $\mathfrak{C}_n$ is the restriction to Φ_n of an ideal product $\mathfrak{A}\mathfrak{B}^{(n)}$ in Φ_ω. Moreover, if B_n is*

of finite rank N as free subgroup of F_n, then $\mathfrak{A} = \mathfrak{A}^{(N)}$ may be assumed. $\mathfrak{C}_n$ is itself a product in Φ_n if, and only if, $\mathfrak{A}^{(N)} = \mathfrak{A}^{(n)}$, and then $\mathfrak{C}_n = \mathfrak{A}_n \mathfrak{B}_n$, that is $C_n = A_n \circ B_n$.

Proof. If $\mathfrak{C}_n = \mathfrak{A}\mathfrak{B}^{(n)} \cap \Phi_n$, then C_n consists, by 9.03, of the components of all elements of the form $(\varepsilon_n \alpha)(\beta \varepsilon_n)$; that is, C_n consists of all "values" of the words in A for "arguments" in B_n. Thus C_n is a word subgroup, that is a fully invariant subgroup of B_n.

Conversely, if C_n is fully invariant in B_n, it is a word subgroup of B_n; there exists, therefore, a set of words $a_i(b_1, b_2, \ldots)$ in the free generators of B_n, such that C_n is the fully invariant closure of this set in B_n. We denote the fully invariant closure in F of the corresponding set of words $a_i(e_1, e_2, \ldots)$ by A. If B_n is of finite rank N, the words a_i are words in $b_1, \ldots, b_N$, hence A is the closure of the words $a_i(e_1, \ldots, e_N)$; therefore the corresponding ideal $\mathfrak{A}$ of Φ_ω satisfies $\mathfrak{A} = \mathfrak{A}^{(N)}$. Moreover, using again 9.03, one confirms easily, that with this choice of $\mathfrak{A}$, $\mathfrak{A}\mathfrak{B}^{(n)} \cap \Phi_n = \mathfrak{C}_n$.

The remaining statements of the theorem now follow from 9.08 and 9.09.

In conclusion, we return to the non-associativity of ideal multiplication in the near-rings Φ_n: we have yet to confirm by an example, that the associative law may, in fact, be violated. The situation is now clear:

If $\mathfrak{A}_n, \mathfrak{B}_n, \mathfrak{C}_n$ are any three ideals in Φ_n, and $\mathfrak{A}^{(n)}, \mathfrak{B}^{(n)}, \mathfrak{C}^{(n)}$ their extensions to Φ_ω, then in Φ_ω, $\mathfrak{A}^{(n)}(\mathfrak{B}^{(n)}\mathfrak{C}^{(n)}) = (\mathfrak{A}^{(n)}\mathfrak{B}^{(n)})\mathfrak{C}^{(n)}$. Also, from 9.07, $\mathfrak{A}^{(n)}\mathfrak{B}^{(n)} \cap \Phi_n = \mathfrak{A}_n \mathfrak{B}_n$ and $\mathfrak{B}^{(n)}\mathfrak{C}^{(n)} \cap \Phi_n = \mathfrak{B}_n \mathfrak{C}_n$, hence $\mathfrak{A}_n(\mathfrak{B}_n \mathfrak{C}_n) = \mathfrak{A}^{(n)}(\mathfrak{B}^{(n)}\mathfrak{C}^{(n)}) \cap \Phi_n = (\mathfrak{A}^{(n)}\mathfrak{B}^{(n)})\mathfrak{C}^{(n)} \cap \Phi_n \supseteq (\mathfrak{A}_n \mathfrak{B}_n)^{(n)} \cdot \mathfrak{C}^{(n)} \cap \Phi_n = (\mathfrak{A}_n \mathfrak{B}_n)\mathfrak{C}_n$. Although $\mathfrak{A}^{(n)}$ and $\mathfrak{B}^{(n)}$ both are generated by elements in Φ_n, their product need not have this property, and in that case the inclusion $(\mathfrak{A}_n \mathfrak{B}_n)^{(n)} \subset \mathfrak{A}^{(n)}\mathfrak{B}^{(n)}$ is proper. If then further the group C_n corresponding to $\mathfrak{C}_n$ is of infinite rank, 9.09 shows that

$$(\mathfrak{A}_n \mathfrak{B}_n)\mathfrak{C}_n \subset \mathfrak{A}_n(\mathfrak{B}_n \mathfrak{C}_n).$$

9.11. Example. We take $n = 2$, and $\mathfrak{A}_2 = \mathfrak{B}_2 = \mathfrak{C}_2$ all as the derived ideal $\mathfrak{D}_2$ of Φ_2. Then $\mathfrak{D}^{(2)}$ is the derived ideal of Φ_ω. Hence: $\mathfrak{D}^{(2)} \cdot \mathfrak{D}^{(2)}$ is the second derived ideal of Φ_ω. (The associated second derived group was denoted by A in Example 2.11, which we shall use here.) $\mathfrak{D}_2\mathfrak{D}_2$ is the second derived ideal of Φ_2, so that $(\mathfrak{D}_2\mathfrak{D}_2)^{(2)}$ is the ideal of Φ_ω corresponding to the two-variable metabelian laws, which, by 2.12, is properly contained in the second derived ideal of Φ_ω:

$$(\mathfrak{D}_2\mathfrak{D}_2)^{(2)} \subset \mathfrak{D}^{(2)}\mathfrak{D}^{(2)}.$$

Since, moreover, the derived group of F_2 is of infinite rank, we have:

9.12. *In Φ_2,* $(\mathfrak{D}_2\mathfrak{D}_2)\mathfrak{D}_2 \subset \mathfrak{D}_2(\mathfrak{D}_2\mathfrak{D}_2)$,

showing that ideal multiplication in Φ_2 is not power associative.

The fully invariant subgroup of F_2 corresponding to the right-hand side of 9.12 is simply the third derived group—obtained as the derived group of the second derived group. The group corresponding to the left-hand side of 9.12

is the two-variable second derived group of the derived group, properly contained in the third derived group—the latter this time thought of as the whole second derived group of the derived group.

10. Appendix: Ideal bases

The elements α_i form a basis for the ideal $\mathfrak{A}$ in Φ_n or Φ_ω, if, and only if, their components form a basis for the corresponding fully invariant subgroup A in F_n or F respectively. In particular, $\mathfrak{A}$ is finitely based if, and only if, A is finitely based. In the free group F of infinite rank, a single word is sufficient to generate A, whenever A is finitely based. The corresponding ideal, generated by a single element, may be reasonably called a principal ideal.

Problem. *Is the product of two principal ideals of Φ_ω necessarily finitely based, that is, itself a principal ideal?*

An infinite ideal basis for such a product can easily be given: If A is the closure in F of the word $a(e_1, \ldots, e_m)$, and B is the closure in F of the word $b(e_1, \ldots, e_n)$, we put

$$a_r = a(e_1 + \cdots + e_r, e_{r+1} + \cdots + e_{2r}, \ldots, e_{(m-1)r+1} + \cdots + e_{mr}),$$

and we take β_r as the "matrix"

$$\beta_r = \begin{pmatrix} b(e_1, \ldots, e_n) \\ b(e_{n+1}, \ldots, e_{2n}) \\ \cdots\cdots\cdots\cdots \\ b(e_{(mr-1)n+1}, \ldots, e_{mrn}) \end{pmatrix},$$

all further components being zero. Then:

10.01. *The elements $c_r = a_r \beta_r$, $r = 1, 2, \ldots$, form a basis for $A \circ B$. Every finite subset of this basis is redundant, as c_r is obtained from c_{r+1} by substituting zeros for some of the generators e_i.*

In special cases, this infinite basis can be reduced to a single element. We give two examples:

10.02. Example. $A = \{2e_1\}$, $B = \{2e_1\} = A$; that is, $A \circ B = A \circ A$ is the "square" of the word group generated by all doubles of elements of F. Here $c_r = 2(2e_1 + \cdots + 2e_r)$. Using that every commutator is a sum of doubles, one shows inductively: $c_r \in \{c_{r-1}\}$ for $r \geq 3$. Therefore

$A \circ A$ is the closure of the single word $2(2e_1 + 2e_2)$.

10.03. Example. $A = \{2e_1\}$, $B = \{3e_1\}$. Here the basis 10.01 for $A \circ B$ consists of the elements $c_r = 2(3e_1 + \cdots + 3e_r)$, $r = 1, 2, \ldots$, and the basis 10.01 for $B \circ A$ consists of the elements $d_r = 3(2e_1 + \cdots + 2e_r)$, $r = 1, 2, \ldots$. In the first case, an argument similar to the one used in Example 10.02 shows that the single word $2(3e_1 + 3e_2)$ is sufficient; in the second case, GRAHAM HIGMAN has shown[10]) that there exists an integer k such that d_r with $r > k$ follows from d_k. Thus:

[10]) Unpublished.

$A \circ B$ is the closure of $2(3e_1+3e_2)$, $B \circ A$ is the closure of $3(2e_1+\cdots+2e_k)$ for some integer k.

In connection with this last example, HIGMAN conjectured that $A \circ B$ is finitely based, whenever A contains some term of the lower central series of F.

Finally, these examples suggest a further question; in § 2 we denoted by $k = k(A)$ the least integer such that A is the closure of a set of words involving at most k variables.

Problem. *Is $k(A \circ B)$ finite, provided $k(A)$ and $k(B)$ are finite? And if the answer is positive, can one give a bound for $k(A \circ B)$ in terms of $k(A)$ and $k(B)$?*

Bibliography

[1] BAER, R.: Groups with proper isomorphic quotient groups. Bull. Amer. Math. Soc. **50**, 267—278 (1944). — [2] BLACKETT, D. W.: Simple and semi-simple near-rings. Proc. Amer. Math. Soc. **4**, 772—785 (1953). — [3] DESKINS, W. E.: A radical for near-rings. Proc. Amer. Math. Soc. **5**, 825—827 (1954). — [4] ETHERINGTON, I. M. H.: Non-associative arithmetics. Proc. Roy. Soc. Edinburgh, Sect. A **62**, 442—453 (1949). — [5] GRUENBERG, K.: Unpublished. — [6] HALL, P.: The splitting properties of relatively free groups. Proc. London Math. Soc. (3) **4**, 343—356 (1954). — [7] LEVI, F.: Über die Untergruppen der freien Gruppen. 2. Mitteilung. Math. Z. **37**, 90—97 (1933). — [8] LEVI, F., and B. L. VAN DER WAERDEN: Über eine besondere Klasse von Gruppen. Abh. math. Sem. Hans. Univ. **9**, 154—158 (1932). — [9] LYNDON, R. C.: Two notes on nilpotent groups. Proc. Amer. Math. Soc. **3**, 579—583 (1952). — [10] MAGNUS, W.: Beziehungen zwischen Gruppen und Idealen in einem speziellen Ring. Math. Ann. **111**, 259—280 (1935). — [11] NEUMANN, B. H.: Identical relations in groups. I. Math. Ann. **114**, 506—525 (1937). — [12] NEUMANN, B. H.: On a problem of Hopf. J. London Math. Soc. **28**, 351—353 (1953). — [13] ROBINSON, A.: On non-associative systems. Proc. Edinburgh Math. Soc. (2) **8**, 111—119 (1949). — [14] WEVER, F.: Über Regeln in Gruppen. Math. Ann. **122**, 334—339 (1950). — [15] ZASSENHAUS, H.: Gruppentheorie, pp. 71—74. 1937.

Department of Mathematics, The University, Hull (England)

(Eingegangen am 9. Oktober 1955)

FOSTER, A. L.
Math. Zeitschr. Bd. 65, S. 70—75 (1956)

Ideals and their Structure in Classes of Operational Algebras

By

ALFRED L. FOSTER

1. Introduction

In the realm of Boolean algebras the classical ideal structure theorem (STONE [1])[1]) asserts: *In a Boolean algebra each (non-unit) ideal is the intersection of its prime ideal — or equivalently, of its maximal ideal — divisors.*

Intimately related to this stands the fundamental subdirect structure theorem: *Each Boolean algebra is isomorphic with a subdirect power of the 2-element Boolean algebra.* A parallel subdirect structure is known for p-rings (MCCOY and MONTGOMERY [2], FOSTER [7]) and again for Post algebras (ROSENBLOOM [4], WADE [5], FOSTER [7]).

On a considerably broader front, in a series of recent papers dedicated to the development of a general "Boolean" theory of operational (or universal) algebras (FOSTER [7] to [12]), it was shown that the entire range of classical Boolean theory — e.g., duality principle, associated "logic", normal expansion theorem (for Boolean functions), structure theory, etc. — may be lifted, in generalized form, to a variety of successively more comprehensive universal algebra levels. Viewed from such a vantage point the Boolean-, Post- and p-ring theories, in particular, emerge as quite special instances of the extended theory.

In the present brief communication our concern is essentially limited to structure theory on the universal algebra level; specifically, we shall transform the two structure theorems of [7] and [8] — both of which are "subdirect" results — to "maximal-ideal" form, corresponding to (and generalizing) the classical Boolean structure theorem of STONE.

2. Some basic notions

In this section we recall several pertinent basic concepts and results of [7], [8]. Let $\mathfrak{A} = (A, o_1, o_2, \ldots)$ be a universal algebra of species $Sp = (n_1, n_2, \ldots)$, that is, an algebra whose basic operations o_i are n_i-ary ($i = 1, 2, \ldots$). $\mathfrak{A}$ is called *functionally (strictly) complete* (f.s.c.) if for each $k = 1, 2, \ldots$ each set-theoretic mapping $f(\xi_1, \xi_2, \ldots, \xi_k)$ of the class $A \times A \times \cdots$ (k^{th} Cartesian power) into the set A is representable in at least one way as a *(strict) $\mathfrak{A}$-function*, i.e., as some composition, via the primitive operations o_i, of indeterminates

[1]) Square brackets refer to the appended bibliography.

$\xi_1, \ldots, \xi_k$ over A. The one-element algebra (of any species), denoted by $\overset{\circ}{|}$, is of course f.s.c. *A primal* algebra (of given species) is defined as a f.s.c. algebra $\neq \overset{\circ}{|}$. Among the properties of such algebras are the following, selected from [7], [8]:

Theorem 1. Let $\mathfrak{P} = (P, \mathfrak{o}, \ldots)$ *be a primal algebra of species* $Sp = (n_1, n_2, \ldots)$. *Then:*

(1) *There exists at least one* n_i *of* Sp *such that* $n_i \geq 1$.

(2) $\mathfrak{P}$ *is a finite algebra (i.e.,* P *is a finite class).*

(3) $\mathfrak{P}$ *is simple (i.e., possesses no proper ideals).*

(4) $\mathfrak{P}$ *possesses no subalgebra (except* $\mathfrak{P}$ *itself).*

(5) *The class of strict identities of* $\mathfrak{P}$ *is a maximal class*[2]).

(6) *The class of strict identities of* $\mathfrak{P}$ *possesses a finite basis*[3]).

For each species [subject to (1) of theorem 1] and for each integer $n > 1$, there exists at least one primal algebra of n elements (= *order* n). Illustrative of the profusion of such algebras we recall the following classes of primals ([7], [8]):

(a) $(F_2, \times, ^\frown)$, the 2-element Boolean algebra ($\times$ = intersection, $^\frown$ = complement).

More generally

(b) $(F_p, \times, ^\frown)$, the prime field, where: $\xi \times \eta$ is the product, and $\xi^\frown = 1 + \xi$ in the prime field $(F_p, \times, +)$.

(c) $(\mathcal{F}_n, \times, ^\frown)$, the "$n$-field", where: $\mathcal{F}_n = \{0, 1, \alpha, \alpha^2, \ldots, \alpha^{n-2}\}$, $\mathcal{F}_n' = \mathcal{F}_n$ less 0, $(\mathcal{F}_n', \times)$ = cyclic group of order $n-1$ $(\alpha^{n-1} = 1)$,

$$0 \times \xi = \xi \times 0 = 0 \,(\xi \in \mathcal{F}),\ \xi^\frown = \text{def}: 0^\frown = 1,\ 1^\frown = \alpha,\ \alpha^\frown = \alpha^2, \ldots, (\alpha^{n-2})^\frown = 0.$$

If the (arbitrary) integer n is chosen as a prime power we have to do with a simple form of the corresponding Galois field. More generally,

(d) $(G_n, \times, ^\frown)$, an "ordered group with null", where: $G_n = \{0, 1, \alpha, \beta, \ldots\}$, $G_n' = G_n$ less 0, $(G_n', \times)$ = arbitrary group of order $n-1$ (with 1 as its identity), $\xi^\frown = \text{def}$ = arbitrary cyclic permutation of the elements of G_n, $0 \times \xi = \xi \times 0 = 0\,(\xi \in G_n)$.

(e) $(P_n, \times, ^\frown)$, the basic Post-algebra of order n, where:

$$P_n = \{0, \alpha_{n-2}, \alpha_{n-1}, \ldots, \alpha_2, \alpha_1, 1\}, \qquad \xi \times \eta = \min(\xi, \eta)$$

where "min" refers to the above sequence as "ascending", $\xi^\frown = \text{def}$:

$$0^\frown = 1,\ 1^\frown = \alpha_1,\ \alpha_1^\frown = \alpha_2, \ldots, (\alpha_{n-2})^\frown = 0.$$

Additional classes of primal algebras are found in [7], [8].

[2]) A strict identity $f(\xi, \ldots) = g(\xi, \ldots)$ of an algebra $\mathfrak{P}$ is one in which f and g are both strict $\mathfrak{P}$-functions. "Maximal" here means: from the strict identities of $\mathfrak{P}$, augmented by any single strict equation (of species Sp) which is not an identity of $\mathfrak{P}$, any desired equation (of species Sp) is logically deducible.

[3]) This finite basis result, [6], while implicit in [7], was first explicitly proved by YAQUB, [13].

3. $\mathfrak{P}$-*algebras; sub-direct structure*

Let $\mathfrak{P}$ be an arbitrary but fixed primal algebra, and let Sp denote its species. An algebra $\mathfrak{A}$ of species Sp is called a $\mathfrak{P}$-*algebra*, or an algebra with *kernel* $\mathfrak{P}$, if $\mathfrak{A}$ satisfies all strict identities of $\mathfrak{P}$ (see earlier footnote). Equivalently stated, the class of $\mathfrak{P}$-algebras may be defined as the class of algebras, of species Sp, which satisfy a certain finite list $I_1, I_2, \ldots, I_t$ of strict equations; here $I_1, \ldots, I_t$ may be chosen as any finite basis for the strict identities of the kernel, $\mathfrak{P}$ [compare with (6) of theorem 1].

In particular the class of algebras with: (a), kernel $(F_2, \times, ^{\circ})$ is identified as the class of Boolean algebras; (b) kernel $(F_p, \times, ^{\circ})$ is coextensive with the class of p-algebras (p-rings) (McCOY and MONTGOMERY [2], FOSTER [11], [12]; ...; (e), kernel $(P_n, \times, ^{\circ})$ is identical with the class of Post algebras of period n (ROSENBLOOM [4], WADE [5], FOSTER [7], [8], etc).

We shall require the following easy

Lemma 2. A homomorphic image of a $\mathfrak{P}$-*algebra is a* $\mathfrak{P}$-*algebra.*

Proof. Let $\mathfrak{A}$ be a $\mathfrak{P}$-algebra and $\mathfrak{B}$ a homomorphic image of $\mathfrak{A}$. Since the strict identities of any operational algebra are inherited by each of its homomorphic images, $\mathfrak{B}$ satisfies all strict identities of $\mathfrak{A}$ and therefore also of $\mathfrak{P}$. The lemma thus follows[4]).

We shall also need a section of the principal structure theorem proved in [7], namely

Theorem 3. Let $\mathfrak{P}$ *be an arbitrary primal algebra. Then the three classes of algebras below are pairwise coextensive (up to isomorphisms):*

(i) *The class of all* $\mathfrak{P}$-*algebras.*

(ii) *The class of all subdirect powers of* $\mathfrak{P}$.

(iii) *The class of all subalgebras of direct powers of* $\mathfrak{P}$.

Further, each finite $\mathfrak{P}$-*algebra is isomorphic with a direct power of* $\mathfrak{P}$.

4. *Primal- and maximal-ideals; ideal structure*

If $\mathfrak{B}$ is a universal algebra and $\mathfrak{I}$ an ideal of $\mathfrak{B}$, $\mathfrak{I}$ is called a *primal-ideal* if the factor algebra $\mathfrak{B}/\mathfrak{I}$ is a primal algebra.

Theorem 4. Let $\mathfrak{P}$ *be a primal algebra and* $\mathfrak{A}$ *a* $\mathfrak{P}$-*algebra.*

(i) *If* $\mathfrak{I}$ *is a primal ideal of* $\mathfrak{A}$ *then* $\mathfrak{A}/\mathfrak{I} \cong \mathfrak{P}$.

(ii) *A primal ideal of* $\mathfrak{A}$ *is a maximal ideal, and conversely.*

(iii) *If* $\mathfrak{A} \neq \overset{\circ}{|}$ (= *one-element algebra*), *then* $\mathfrak{A}$ *possesses at least one primal ideal.*

Proof. (i) We shall assume $\mathfrak{A} \neq \overset{\circ}{|}$ since in this case the theorem is vacuously true. Let $\mathfrak{I}$ be a primal ideal, whence $\mathfrak{A}/\mathfrak{I} \cong \mathfrak{Q} =$ primal algebra. Again $\mathfrak{Q}$, as homomorphic image of $\mathfrak{A}$, is a $\mathfrak{P}$-algebra (lemma 2). By theorem 3 the (finite) algebra $\mathfrak{Q}$ is then isomorphic with a direct power of $\mathfrak{P}$. This power could only be the first, i.e. $\mathfrak{Q} \cong \mathfrak{P}$, since otherwise the primal $\mathfrak{Q}$ would possess a proper ideal [(3) of theorem 1]. This proves (i).

[4]) Clearly the lemma does not require $\mathfrak{P}$ to be primal.

(ii) Let $\mathfrak{I}$ be a primal ideal of $\mathfrak{A}$. Then $\mathfrak{A}/\mathfrak{I} \cong \mathfrak{P}$, by part (i), and therefore $\mathfrak{I}$ is maximal; for if not, $\mathfrak{P}$ would possess a proper homomorphic image, contrary to (3) of theorem 1. Conversely, suppose $\mathfrak{I}$ is maximal. Then $\mathfrak{A}/\mathfrak{I}$, as homomorphic image of $\mathfrak{A}$, is a $\mathfrak{P}$-algebra (lemma 2), and $\mathfrak{A}/\mathfrak{I}$ is isomorphic with a subdirect power of $\mathfrak{P}$ (theorem 3). This subdirect power must be $\mathfrak{P}$ itself - whence $\mathfrak{A}/\mathfrak{I} \cong \mathfrak{P}$ and $\mathfrak{I}$ is a primal ideal —, for otherwise $\mathfrak{A}/\mathfrak{I}$ would possess a proper homomorphic image, contradicting the assumption that $\mathfrak{I}$ is maximal.

(iii) If $\mathfrak{A}$ is a primal algebra then 0, the null ideal of $\mathfrak{A}$, is a primal ideal since $\mathfrak{A}/0 \cong \mathfrak{A}$. Quite generally (theorem 3), $\mathfrak{A}$ is isomorphic with a subdirect power of $\mathfrak{P}$,

$$\alpha \in \mathfrak{A},\ \alpha = (\alpha_1, \ldots, \alpha_i, \ldots);\ \alpha_1 \in \mathfrak{P}, \ldots, \alpha_i \in \mathfrak{P}, \ldots$$

whence each ideal $\mathfrak{I}_i$, corresponding to the "natural" homomorphism $\alpha \to \alpha_i$, is clearly a primal ideal since $\mathfrak{A}/\mathfrak{I}_i \cong \mathfrak{P}$.

The first ideal-structure theorem is now given by

Theorem 5. Let $\mathfrak{P}$ be a primal algebra and $\mathfrak{A}$ a $\mathfrak{P}$-algebra. Then each non-unit ideal of $\mathfrak{A}$ is the intersection of its maximal (= primal) ideal divisors.

Proof. We again assume $\mathfrak{A} \neq \overset{\circ}{|}$ since in this case the theorem is vacuously true. As before let 0 be the null ideal of $\mathfrak{A}$, whence $\mathfrak{A}/0 \cong \mathfrak{A}$. We shall need the following theorem of BIRKHOFF [6], — see also MCCOY [3]:

Theorem 6. Let $\{\ldots, \mathfrak{G}_i, \ldots\}$ be a class of universal algebras $\mathfrak{G}_i$, all of the same species. A necessary and sufficient condition for an algebra $\mathfrak{B}$ to be isomorphic with a subdirect product of the algebras $\mathfrak{G}_i$ is that $\mathfrak{B}$ possess a class $\{\ldots, \mathfrak{I}_i, \ldots\}$ of ideals, $\mathfrak{I}_i$, (1°) whose intersection is the null-ideal,

$$\cap\{\ldots, \mathfrak{I}_i, \ldots\} = 0,$$

and such that (2°) for the corresponding factor algebras,

$$\mathfrak{B}/\mathfrak{I}_i \cong \mathfrak{G}_i \quad \textit{(for each index } i\textit{)}.$$

We return to the proof of theorem 5. Using one half of theorem 6 — with all $\mathfrak{G}_i$ taken as the same primal algebra, $\mathfrak{P}$ — and with due regard to theorem 3, it follows that: $\mathfrak{A}$ (or what is the same, the null ideal of $\mathfrak{A}$) is the intersection of a class $\{\ldots, \mathfrak{I}_i, \ldots\}$ of ideals $\mathfrak{I}_i$ such that

$$\mathfrak{A}/\mathfrak{I}_i \cong \mathfrak{P} \qquad \text{(for each index } i\text{)}.$$

It follows that these ideals $\mathfrak{I}_i$ are primal, and therefore maximal in $\mathfrak{A}$ (theorem 4).

Let now $\mathfrak{I}$ be an arbitrary ideal in $\mathfrak{A}$, and consider the factor algebra

$$\mathfrak{B} = \text{def} = \mathfrak{A}/\mathfrak{I}.$$

By lemma 2, $\mathfrak{B}$ is again a $\mathfrak{P}$-algebra and therefore by the foregoing conclusion (with $\mathfrak{B}$ replacing $\mathfrak{A}$) we have: $\mathfrak{B}$ (or the null ideal of $\mathfrak{B}$) is expressible as the intersection of a class of maximal ideals in $\mathfrak{B}$. Now, in an obvious way, to each ideal of the factor algebra $\mathfrak{B}$ corresponds an ideal in $\mathfrak{A}$, and in this correspondence: (1) to the null ideal of $\mathfrak{B}$ corresponds the ideal $\mathfrak{I}$ of $\mathfrak{A}$,

and (2) to a maximal ideal of $\mathfrak{B}$ corresponds a maximal ideal of $\mathfrak{A}$. Hence the ideal $\mathfrak{I}$ of $\mathfrak{A}$ is expressible as the intersection of a class of maximal ideals of $\mathfrak{A}$. Thus $\mathfrak{I}$ is the intersection of all its maximal ideal divisors, and theorem 5 is proved.

5. *The second ideal structure theorem*

The concept $\mathfrak{P}$-algebra is perforce based on a single (primal) algebra as kernel. In [8] a multiple-kernel theory was initiated, where such a kernel now consists of an appropriately selected class of primal algebras. We shall briefly recall a pertinent segment of this theory.

Algebras $\mathfrak{A}_1, \mathfrak{A}_2, \ldots, \mathfrak{A}_k$, all of the same species Sp,

$$\mathfrak{A}_i = (A_i, o, \ldots)$$

are said to be *strictly independent* if, corresponding to each given set (sequence) of k expressions of species Sp,

$$\Phi_1(\xi, \ldots),\ \Phi_2, \ldots, \Phi_k$$

there exists at least one fixed expression $\chi(\xi, \ldots)$ such that

$$\Phi_i(\xi, \ldots) = \chi(\xi, \ldots) \pmod{\mathfrak{A}_i}, \qquad (i = 1, 2, \ldots, k). \tag{5.1}$$

Here an *expression* is simply one or more indeterminante symbols $\xi, \ldots$ combined via the primitive operations $o, \ldots$ of Sp; in each algebra $\mathfrak{A}$ of species Sp, a given expression represents a strict $\mathfrak{A}$-function. Further, (5.1) is to be read: $\Phi_i = \chi$ is a (strict) identity of the algebra $\mathfrak{A}_i$.

Let $\mathfrak{P}_1, \mathfrak{P}_2, \ldots, \mathfrak{P}_k$ $(k \geq 1)$ be a finite set of independent primal algebras, all of a given species Sp. An algebra $\mathfrak{A}$ of species Sp is said to be

Def 1. A $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$-*algebra* if $\mathfrak{A}$ satisfies all strict identities common to $\mathfrak{P}_1, \ldots, \mathfrak{P}_k$, — or, equivalently, if $\mathfrak{A}$ satisfies all strict identities of the algebra $\mathfrak{P}_1 \times \mathfrak{P}_2 \times \cdots \times \mathfrak{P}_k$ (direct product).

Def 2. *Subdirectly representable* in $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$ if $\mathfrak{A}$ is isomorphic with a suitable subdirect product of suitable subdirect powers of $\mathfrak{P}_1, \ldots, \mathfrak{P}_k$.

Def 3. *Semi-directly representable* in $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$ if $\mathfrak{A}$ is isomorphic with the direct product of suitable subdirect powers of $\mathfrak{P}_1, \ldots, \mathfrak{P}_k$.

Def 4. *Directly representable* in $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$ if $\mathfrak{A}$ is isomorphic with the direct product of suitable direct powers of $\mathfrak{P}_1, \ldots, \mathfrak{P}_k$.

A portion of the fundamental structure theorem of [8] asserts[5])

Theorem 7. *Let $\mathfrak{P}_1, \ldots, \mathfrak{P}_k$ be an arbitrary independent class of primal algebras. Then the following four classes of algebras are pairwise coextensive up to isomorphisms:*

(i) *The class of all $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$-algebras.*

(ii) *The class of all algebras subdirectly representable in $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$.*

(iii) *The class of all algebras semi-directly representable in $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$.*

[5]) A second part asserts a certain *unicity* for the subdirect representation. The transformation of this unicity into primal-ideal form is not here considered.

(iv) *The class of all subalgebras of algebras directly representable in* $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$.

Further, a finite $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$ *- algebra is directly representable in* $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$.

If in the proofs of theorems 4 and 5 we now appeal to theorem 7 rather than theorem 3, arguments parallel to those already given yield

Theorem 8. Let $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$ *be a strictly independent class of primal algebras, and let* $\mathfrak{A}$ *be a* $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$*-algebra. Then*

(i) *If* $\mathfrak{I}$ *is a primal ideal of* $\mathfrak{A}$, $\mathfrak{A}/\mathfrak{I} \cong \mathfrak{P}_i$, *where* $\mathfrak{P}_i$ *is a primal belonging to the class* $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$.

(ii) *A primal ideal of* $\mathfrak{A}$ *is a maximal ideal, and conversely.*

(iii) *If* $\mathfrak{A} \neq \overset{\circ}{|}$ *then* $\mathfrak{A}$ *possesses at least one primal ideal.*

Theorem 9. Let $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$ *be a strictly independent class of primal algebras, and let* $\mathfrak{A}$ *be a* $\{\mathfrak{P}_1, \ldots, \mathfrak{P}_k\}$*-algebra. Then each non-unit ideal of* $\mathfrak{A}$ *is the intersection of its maximal* (= *primal*) *ideal divisors.*

The applications of theorems 8, 9 are widespread. Thus for instance (see § 2),

from example (b): if $p_1, p_2, \ldots, p_k$ are distinct primes, the prime-fields $\{(F_{p_1}, \times, \hat{}), \ldots, (F_{p_k}, \times, \hat{})\}$ form an independent primal class;

from example (c): if $n_1, n_2, \ldots, n_k$ are any distinct integers, with each $n_i > 1$, the corresponding n-fields $\{(\mathcal{F}_{n_1}, \times, \hat{}), \ldots, (\mathcal{F}_{n_k}, \times, \hat{})\}$ form an independent primal class;

from example (e): if $n_1, n_2, \ldots, n_k$ are any distinct integers, with each $n_i > 1$, the corresponding basic Post-algebras $\{(P_{n_1}, \times, \hat{}), \ldots, (P_{n_k}, \times, \hat{})\}$ form an independent primal class.

(For these and other strictly independent primal classes see [7].)

Bibliography

[1] Stone, M. H.: The theory of representations of Boolean algebras. Trans. Amer. Math. Soc. **40**, 37—111 (1936). — [2] McCoy, N. H., and Deane Montgomery: A representation of generalized Boolean rings. Duke Math. J. **3**, 455—459 (1937. — [3] McCoy, N. H.: Subdirect sums of rings. Bull. Amer. Math. Soc. **53**, 856—877 (1947). — [4] Rosenbloom, P.: The Elements of Mathematical Logic. The Dover Series (1950). — [5] Wade, L. I.: Post algebras and rings. Duke Math. J. **12**, 389—395 (1945). — [6] Birkhoff, G.: Subdirect unions in universal algebras. Bull. Amer. Math. Soc. **50**, 764—768(1944). — [7] Foster, A. L.: Generalized "Boolean" theory of universal algebras, Part I: Subdirect sums and normal representation theorem. Math. Z. **58**, 306—336 (1953), and Part II: Identities and subdirect sums of functionally complete algebras. Math. Z. **59**, 191—199 (1953). — [8] Foster, A. L.: The identities of—and unique subdirect factorization within—classes of universal algebras. Math. Z. **62**, 171—188 (1955). — [9] Foster, A. L.: The idempotent elements of a commutative ring form a Boolean algebra; ring duality and transformation theory. Duke Math. J. **12**, 143—152 (1945). — [10] Foster, A. L.: On n-ality theories in rings and their logical algebras, etc. Amer. J. Math. **72**, 101—123 (1950). — [11] Foster, A. L.: p-Rings and their Boolean-vector representation. Acta Math. Stockh. **84**, 231—261 (1950). — [12] Foster, A. L.: p-Rings and ring logics. Univ. California Publ. Math. **1**, 385—396 (1951). — [13] Yaqub, A.: On the theory of ring-logics and universal algebras. Doctoral dissertation, Univ. of California, Berkeley, (1955).

University of California, Berkeley

(Eingegangen am 24. Oktober 1955)

KRULL, W.
Math. Zeitschr. Bd. 65, S. 76—90 (1956)

Über reelle Radikalkörper

Von
WOLFGANG KRULL

Die hier behandelten Probleme gehören zwar ihrer Natur nach in den Rahmen der „elementaren" bzw. „klassischen" Algebra, sie scheinen mir aber aus den folgenden Gründen bemerkenswert:

1. Während man mit Hilfe der Galoisschen Theorie die Radikalkörper schlechthin in allgemein bekannter Weise völlig befriedigend charakterisieren kann, findet sich über *reelle Radikalkörper* (in Zukunft kurz r. R.-Körper)[1]) in der algebraischen Lehrbuchliteratur im allgemeinen nicht viel mehr als der folgende *p-Konjugiertensatz:* Ist der Grad des r. R.-Körpers $\mathfrak{K}$ über seinem Grundkörper $\mathfrak{K}_0$ eine Primzahl $p \neq 2$, so sind alle echten Konjugiertenkörper von $\mathfrak{K}$ über $\mathfrak{K}_0$ komplex[2]). Indessen ist es, wie in § 2 gezeigt wird, leicht möglich, die Klasse der r. R.-Körper von Primzahlgrad genau zu charakterisieren, und man stößt dabei auf zwei bemerkenswerte, meines Wissens bisher nicht beachtete Tatsachen[3]):

a) Die Gaußschen Primzahlen, also die von der Form $2^m + 1$, nehmen eine Sonderstellung ein. Nur bei ihnen liegen die Verhältnisse so einfach, wie man bei einer voreiligen Verallgemeinerung der bei den Körpern dritten Grades gemachten Erfahrungen allgemein erwarten könnte.

b) Will man bei einer nicht-Gaußschen Primzahl p eine hinreichende Kennzeichnung für die r. R.-Körper p-ten Grades erhalten, so muß man eine numerische Invariante heranziehen, die unter zahlentheoretischen Gesichtspunkten von KUMMER bei der Entwicklung der Theorie der nach ihm benannten Körper eingeführt wurde. Das heißt aber, man muß in gewissem Sinne den

[1]) Unter einem reellen Radikalkörper (über dem Grundkörper $\mathfrak{K}_0$) verstehen wir einen Körper, der aus dem im üblichen Sinne reellen Körper $\mathfrak{K}_0$ durch Adjunktion einer reellen Radikalgröße entsteht, wobei die reelle Radikalgröße ihrerseits dadurch ausgezeichnet ist, daß zu ihrer Bildung außer den Elementen von $\mathfrak{K}_0$ nur *reelle* Nullstellen *reeller* binomischer Gleichungen benutzt werden. Vgl. im übrigen die exakten Definitionen am Anfang von § 1.

[2]) Ein echt konjugierter Körper von $\mathfrak{K}$ ist natürlich ein zu $\mathfrak{K}$ (über $\mathfrak{K}_0$) konjugierter Körper $\mathfrak{K}' \neq \mathfrak{K}$. Gewöhnlich erscheint der p-Konjugiertensatz in der Form: Ist eine Nullstelle α des über dem reellen Körper $\mathfrak{K}_0$ irreduziblen Polynoms $P(x)$ vom Primzahlgrade $p \neq 2$ eine reelle Radikalgröße, so sind alle anderen Nullstellen von $p(x)$ komplex. — Aber es zeigt sich bei weitergehenden Untersuchungen, daß man in der Theorie der r. R.-Körper zweckmäßig nur mit Körpern und ihren Isomorphismen und Automorphismen, nicht aber mit den erzeugenden Polynomen und definierenden Gleichungen der Körper arbeitet.

[3]) Den Inhalt von § 2 habe ich in den letzten 6 Jahren öfters in Vorlesungen und Seminaren behandelt, da sich hier eine ausgezeichnete Möglichkeit bietet, den Anfänger an einem historisch interessanten nichttrivialem Problem in die Anwendung der Galoisschen Theorie einzuführen.

Rahmen der klassischen Galoisschen Theorie überschreiten, — bei einem Problem, das letzten Endes vom „casus irreducibilis" der Gleichungen dritten Grades, also aus dem Gedankenkreis CARDANOs herstammt.

2. Es liegt äußerst nahe zu fragen: Gibt es eine sinnvolle Verallgemeinerung des fundamentalen p-Konjugiertensatzes auf reelle Radikalkörper beliebigen Grades[4])? Hier ist das wichtigste Ergebnis die Erkenntnis, daß nur durch eine ganz neue Begriffsbildung eine befriedigende Antwort gewonnen werden kann. Die Untersuchung beliebiger r.R.-Körper ist vor allem deshalb schwierig, weil diese Körper in der Regel *nicht* über ihrem Grundkörper $\mathfrak{R}_0$ normal sind. Um trotzdem vorwärts zu kommen, muß man beachten, daß jeder über $\mathfrak{R}_0$ normale Körper $\mathfrak{N}$ einen eindeutig bestimmten (mengentheoretisch), *maximalen* r.R.-Körper enthält. Beschränkt man sich auf maximale r.R.-Körper, so kommt man zu einfachen und in ihrer präzisen Form befriedigenden Sätzen, zu denen es, wie durch Beispiele gezeigt werden kann, bei beliebigen r.R.-Körpern kein Gegenstück gibt. Zunächst zeigt man leicht: Zu jedem maximalen r.R.-Körper $\mathfrak{R}$ existiert eine Körperkette

$$\mathfrak{R}_0 \subset \mathfrak{R}_1 \subset \cdots \subset \mathfrak{R}_m = \mathfrak{R},$$

bei der die Relativgrade $[\mathfrak{R}_i : \mathfrak{R}_{i-1}]$ sämtlich Primzahlen sind. Mit Hilfe dieses Satzes und der Theorie der r.R.-Körper vom Primzahlgrad gelingt dann weiter durch einen etwas mühsamen Induktionsschluß die gewünschte Verallgemeinerung des p-Konjugiertensatzes in der folgenden Form: Die zu einem maximalen r.R.-Körper echt konjugierten Körper sind alle komplex. — Die Anzahl der Automorphismen eines maximalen r.R.-Körpers ist stets eine Potenz von 2.

3. Die Erkenntnis von der Sonderstellung der maximalen r.R.-Körper zeigt einen Weg, auf dem vielleicht das — sicher schwierige — Problem der Charakterisierung aller r.R.-Körper ohne Gradbeschränkung erfolgreich angegriffen werden kann. Man wird voraussichtlich in erster Linie zu fragen haben: Welche Eigenschaften muß ein Normalkörper $\mathfrak{N}$ besitzen, wenn er der *kleinste* Normalkörper sein soll, der den in ihm steckenden maximalen r.R.-Körper umfaßt? — (In der vorliegenden Arbeit, in der der Konjugiertensatz für maximale r.R.-Körper das Hauptergebnis ist, gehen wir auf dieses Problem nicht näher ein.)

§1. Grundeigenschaften reeller Radikalkörper

Alle auftretenden Körper sollen in den Gaußschen Körper aller komplexen Zahlen eingebettet sein. Das Wort „reeller Körper" hat also den klassischen Sinn, nicht den formalen von ARTIN und VAN DER WAERDEN[5]). Wir stützen uns auf den folgenden wohlbekannten

[4]) Dabei ist nur an solche Sätze gedacht, die eine ähnlich scharfe Eindeutigkeitsaussage enthalten wie der Konjugiertensatz selbst.

[5]) Allerdings könnte man statt des Körpers aller gewöhnlichen reellen bzw. aller komplexen Zahlen einen beliebigen geordneten, reell abgeschlossenen Körper P bzw. den zugehörigen algebraisch abgeschlossenen Körper P^* wählen, ausschließlich Unter-

Hilfssatz 1. a) *Ist* $x^p - a$ $(a \in \mathfrak{K})$ *ein über dem Körper* $\mathfrak{K}$ *irreduzibles Binom, und wird* $x^p - a$ *über dem Körper* $\mathfrak{L} > \mathfrak{K}$ *reduzibel, so enthält* $\mathfrak{L}$ *mindestens eine Nullstelle von* $x^p - a$, *es zerfällt also speziell* $x^p - a$ *für* $p = 2$ *über* $\mathfrak{L}$ *in Linearfaktoren.* b) *Sind* $\mathfrak{K}$ *und* $\mathfrak{L}$ *reell, so enthält* $\mathfrak{L}$ *im Falle der Reduzibilität von* $x^p - a$ $(p \neq 2)$ *genau eine Nullstelle von* $x^p - a$, *nämlich die einzige reelle.*

Wir nennen $\mathfrak{K}$ einen *eigentlichen R-Körper* (Radikalkörper) über $\mathfrak{K}_0 < \mathfrak{K}$, wenn eine Körperkette $\mathfrak{K}_i = \mathfrak{K}_{i-1}(\alpha)$ $(i = 1, \ldots, m;\ \mathfrak{K}_m = \mathfrak{K})$ existiert, bei der α_i, jeweils Nullstelle eines über $\mathfrak{K}_{i-1}$ irreduziblen Binoms $x^{p_i} - a_i$ $(a_i \in \mathfrak{K}_{i-1})$ vom Primzahlgrade p_i ist. Sind $\mathfrak{K}_0$ und alle α_i, also auch $\mathfrak{K}$, *reell*, so wird $\mathfrak{K}$ als *eigentlicher r.R.-Körper* (reeller Radikalkörper) über $\mathfrak{K}_0$ bezeichnet. Wo einer Untersuchungsreihe ein fester Ausgangskörper $\mathfrak{K}_0$ zugrunde liegt, wird „über $\mathfrak{K}_0$" weggelassen. Ein *R.-Körper* bzw. *r.R.-Körper* (über $\mathfrak{K}_0$) schlechthin ist jeder ($\mathfrak{K}_0$ enthaltende) Unterkörper eines eigentlichen R.-Körpers bzw. eigentlichen r.R.-Körpers. Jedes Element eines R.-Körpers bzw. r.R.-Körpers heißt *Radikalgröße* bzw. *reelle Radikalgröße* über $\mathfrak{K}_0$. Eine *erzeugende Radikalkette* des R.-Körpers $\mathfrak{K}$ über $\mathfrak{K}_0$ ist eine Kette von der Art, wie sie bei der Definition von $\mathfrak{K}$ benutzt wurde.

Satz 1. a) Ist $\mathfrak{K}$ *(eigentlicher) r.R.-Körper über* $\mathfrak{K}_0$, *so ist es* $\mathfrak{K}$ *auch über jedem Körper* $\mathfrak{K}_0'$ *zwischen* $\mathfrak{K}_0$ *und* $\mathfrak{K}$. *b) Der Vereinigungskörper* $\mathfrak{K} \cdot \mathfrak{L}$ *zweier (eigentlicher) r.R.-Körper über* $\mathfrak{K}_0$ *ist selbst ein (eigentlicher) r.R.-Körper über* $\mathfrak{K}_0$. *c) Ist* $\mathfrak{K}$ *ein (eigentlicher) r.R.-Körper über* $\mathfrak{K}_0$, $\mathfrak{L}$ *ein (eigentlicher) r.R.-Körper über* $\mathfrak{K}$, *so ist* $\mathfrak{L}$ *auch ein (eigentlicher) r.R.-Körper über* $\mathfrak{K}_0$.

a) und b) gelten ohne „eigentlich" trivialerweise, wenn sie mit „eigentlich" richtig sind. Um sie für eigentliche r.R.-Körper zu beweisen, sei $\alpha_1, \ldots, \alpha_n$ bzw. $\beta_1, \ldots, \beta_m$ eine erzeugende Radikalkette von $\mathfrak{K}$ bzw. $\mathfrak{L}$ über $\mathfrak{K}_0$. Dann folgt leicht aus Hilfssatz 1, vor allem 1b): a) Nach Weglassung aller der α_i bei denen $x^p - a_i$ über $\mathfrak{K}_0' \cdot \mathfrak{K}(\alpha_1, \ldots, \alpha_{i-1})$ reduzibel wird, wird die übrigbleibende Kette $\alpha_{i_1}, \ldots, \alpha_{i_r}$ eine erzeugende Radikalkette von $\mathfrak{K}$ über $\mathfrak{K}_0'$! b) Läßt man die β_k weg, für die $x^p - b_k$ über $\mathfrak{K} \cdot \mathfrak{K}_0(\beta_1, \ldots, \beta_{k-1})$ reduzibel ist, so stellt die mit den übrigbleibenden β gebildete Kette $\alpha_1, \ldots, \alpha_n, \beta_{k_1}, \ldots, \beta_{k_s}$ eine erzeugende Radikalkette von $\mathfrak{K} \cdot \mathfrak{L}$ über $\mathfrak{K}_0$ dar. c) Behauptung c) ist nach Definition bei *eigentlichen* r.R.-Körpern trivial. Im allgemeinen Fall gibt es sicher einen über $\mathfrak{K}_0$ bzw. $\mathfrak{K}$ eigentlichen r.R.-Körper $\mathfrak{K}^* \supseteq \mathfrak{K}$ bzw. $\mathfrak{L}^* \supseteq \mathfrak{L}$. Nach a) ist $\mathfrak{K}^*$ und nach b) auch $\mathfrak{K}^* \cdot \mathfrak{L}^*$ eigentlicher r.R.-Körper über $\mathfrak{K}$. Nach a) folgt weiter, daß $\mathfrak{K}^* \cdot \mathfrak{L}^*$ auch eigentlicher r.R.-Körper über $\mathfrak{K}^*$ sein muß, und da $\mathfrak{K}^*$ seinerseits eigentlicher r.R.-Körper über $\mathfrak{K}_0$, stellt $\mathfrak{K}^* \cdot \mathfrak{L}^* \supseteq \mathfrak{L}$ einen eigentlichen r.R.-Körper über $\mathfrak{K}_0$ dar.

---

körper von P^* betrachten und die Elemente von P auf Grund der gegebenen Ordnung als reell auszeichnen. Aber diese Verallgemeinerung ist offenbar uninteressant. Der entscheidende Punkt ist der, daß wir — zum mindesten dann, wenn unerträglich langwierige Konjugiertenbetrachtungen vermieden werden sollen — nicht allein von einem reellen, d.h. für uns geordneten, Grundkörper $\mathfrak{K}_0$ ausgehen können. Wir müssen auch von jedem auftretenden Oberkörper $\mathfrak{K}$ von $\mathfrak{K}_0$ von vornherein wissen, ob er reell (also in Fortsetzung der Ordnung von $\mathfrak{K}_0$ geordnet) ist oder nicht.

Um die Bedeutung von Satz 1 zu würdigen, beachte man: *Sind* $\mathfrak{K}$ *und* $\mathfrak{L}$ *eigentliche R-Körper über* $\mathfrak{K}_0$, *so braucht* $\mathfrak{K} \cdot \mathfrak{L}$ *kein eigentlicher R-Körper über* $\mathfrak{K}_0$ *zu sein.* Es sei etwa $\varepsilon = \exp \frac{2\pi i}{7}$ eine siebte Einheitswurzel, $\mathfrak{K}_0$ der Körper der rationalen Zahlen, α und $\alpha' = \varepsilon \cdot \alpha$ Nullstellen von $x^7 - 2$. Dann sind zwar $\mathfrak{K} = \mathfrak{K}_0(\alpha)$ und $\mathfrak{L} = \mathfrak{K}_0(\alpha')$, nicht aber $\mathfrak{K} \cdot \mathfrak{L} = \mathfrak{K}_0(\alpha, \alpha')$ eigentliche R-Körper über $\mathfrak{K}_0$. Denn wir haben $\mathfrak{K} \cdot \mathfrak{L} = \mathfrak{K}_0(\alpha, \varepsilon)$, es ist $\mathfrak{K} \cdot \mathfrak{L} = \mathfrak{N}$ über $\mathfrak{K}_0$ normal, und es gilt die Gradgleichung $[\mathfrak{N} : \mathfrak{K}_0] = 6$. Wäre also $\mathfrak{N}$ ein eigentlicher R-Körper über $\mathfrak{K}_0$, so müßte in einer erzeugenden Radikalkette $\alpha_1, \alpha_2, \alpha_3$ ein α_i auftreten, das über $\mathfrak{K}_0(\alpha_1, \ldots, \alpha_{i-1})$ Nullstelle eines Binoms $x^3 - a_i$ wäre. Wegen seiner Normalität müßte dann $\mathfrak{N}$ auch $\exp \frac{2\pi i}{3}$ enthalten, was nicht zutrifft. — Der Satz, daß $\mathfrak{K} \cdot \mathfrak{L}$ ein R-Körper über $\mathfrak{K}_0$ sein muß, wenn es $\mathfrak{K}$ und $\mathfrak{L}$ sind, ist zwar richtig, aber sein Beweis ist wesentlich umständlicher als der von Satz 1 b). Man muß, da Hilfssatz 1 b) nicht benutzt werden kann, zu mindesten einen neuen Hilfsbegriff einführen, nämlich den der „normierten Radikalketten"[6]). Aus Satz 1 b) folgt sofort: In jedem Körper $\mathfrak{M} > \mathfrak{K}_0$ gibt es einen größten r. R.-Körper $\mathfrak{K}$, nämlich den Vereinigungskörper aller der $\mathfrak{K}' \subseteq \mathfrak{M}$, die r. R.-Körper über $\mathfrak{K}_0$ sind[7]). Wir wollen nun $\mathfrak{K}$ als einen *maximalen r. R.-Körper über* $\mathfrak{K}_0$ bezeichnen, wenn $\mathfrak{K}$ der größte r. R.-Körper in dem zu $\mathfrak{K}$ über $\mathfrak{K}_0$ gehörigen Normalkörper $\mathfrak{N}$ (also in dem kleinsten, $\mathfrak{K}$ umfassenden über $\mathfrak{K}_0$ normalen Körper $\mathfrak{N}$) ist. Neben Satz 1 a) tritt dann ergänzend:

Satz 1 a'). *Ist* $\mathfrak{K}$ *maximaler r. R.-Körper über* $\mathfrak{K}_0$, *so ist es* $\mathfrak{K}$ *auch über jedem Körper* $\mathfrak{K}_0'$ *zwischen* $\mathfrak{K}_0$ *und* $\mathfrak{K}$.

Ist nämlich $\mathfrak{N}$ bzw. $\mathfrak{N}'$ der zu $\mathfrak{K}$ über $\mathfrak{K}_0$ bzw. $\mathfrak{K}_0'$ gehörige Normalkörper, so ist jedenfalls $\mathfrak{N}' \subseteq \mathfrak{N}$. Ist ferner $\mathfrak{L} \subseteq \mathfrak{N}'$ r. R.-Körper über $\mathfrak{K}_0'$, so ist es $\mathfrak{L}$ auch über $\mathfrak{K}_0$ [nach Satz 1 c)], und aus $\mathfrak{L} \subseteq \mathfrak{N}$ folgt $\mathfrak{L} \subseteq \mathfrak{K}$.

Satz 2. *Ist* $\mathfrak{K}$ *ein maximaler r. R.-Körper über* $\mathfrak{K}_0$, *so gibt es stets eine Körperkette* $\mathfrak{K}_0 < \mathfrak{K}_1 < \cdots < \mathfrak{K}_n = \mathfrak{K}$, *bei der* $[\mathfrak{K}_i : \mathfrak{K}_{i-1}] = p_i$ *Primzahl* $(i = 1, \ldots, n)$.

Wegen Satz 1a') genügt es offenbar, zu zeigen: Es gibt einen Körper $\mathfrak{K}_1 \subseteq \mathfrak{K}$ bei dem $[\mathfrak{K}_1 : \mathfrak{K}_0] = p$ Primzahl. Es sei $\mathfrak{N}$ der zu $\mathfrak{K}$ über $\mathfrak{K}_0$ gehörige Normalkörper, $\mathfrak{K}^* \supseteq \mathfrak{K}$ ein eigentlicher r. R.-Körper über $\mathfrak{K}_0$. Ferner sei $\alpha_1, \ldots, \alpha_n$ eine erzeugende Radikalkette von $\mathfrak{K}^*$ über $\mathfrak{K}_0$, und es sei j so bestimmt, daß $\mathfrak{L} \cap \mathfrak{N} = \mathfrak{K}_0$, $\mathfrak{L}(\alpha_j) \cap \mathfrak{N} > \mathfrak{K}_0$ mit $\mathfrak{L} = \mathfrak{K}_0(\alpha_1, \ldots, \alpha_{j-1})$. Dann hat $\mathfrak{L}(\alpha_j) \cap \mathfrak{N} = \mathfrak{K}_1$ über $\mathfrak{K}_0$ den gleichen Primzahlgrad wie $\mathfrak{L}(\alpha_j)$ über $\mathfrak{L}$; ist etwa nämlich $\mathfrak{K}_1 = \mathfrak{K}_0(\beta)$ und $p(x)$ das über $\mathfrak{K}$ irreduzible Polynom mit der Nullstelle β, so muß $p(x)$ wegen der Normalität von $\mathfrak{N}$ und $\mathfrak{L} \cap \mathfrak{N} = \mathfrak{K}_0$ auch über $\mathfrak{L}$ irreduzibel

[6]) Zu den normierten Radikalketten vgl. etwa O. Haupt, Lehrbuch der Algebra, Band II (2. Aufl. 1953). Mit zu schwerem Geschütz schießen, hieße es, wenn man hier den Satz von der gruppentheoretischen Charakterisierung der normalen Radikalkörper heranziehen würde.

[7]) Wir beschränken uns stillschweigend auf endliche Erweiterungen des gegebenen Grundkörpers $\mathfrak{K}_0$, da auch der Begriff der r. R.-Körper nur für endliche Oberkörper von $\mathfrak{K}_0$ eingeführt wurde.

sein. $\mathfrak{K}_1$ ist als Unterkörper von $\mathfrak{K}^*$ ein r.R.-Körper über $\mathfrak{K}_0$ und aus der Maximalität von $\mathfrak{K}$ folgt $\mathfrak{K}_1 \subseteq \mathfrak{K}$.

Daß bei Satz 2 auf die Maximalität von $\mathfrak{K}$ über $\mathfrak{K}_0$ nicht verzichtet werden kann, zeigt das Beispiel eines Polynoms vierten Grades $p(x) = (x-\alpha_1) \cdot (x-\alpha_2) \cdot (x-\alpha_3) \cdot (x-\alpha_4)$ über dem Körper $\mathfrak{K}_0$ der rationalen Zahlen mit zwei reellen und zwei konjugiert komplexen Nullstellen α_1, α_2 bzw. α_3, α_4, dessen Gruppe die symmetrische ist. Hier hat die übliche kubische Resolvente eine reelle Nullstelle $\gamma_1 = (\alpha_1+\alpha_2-\alpha_3-\alpha_4)^2 = \beta_1^2$ $(\beta_1 > 0)$ und zwei konjugiert komplexe Nullstellen γ_2, γ_3. Nach CARDANO ist also $\mathfrak{K}_0(\gamma_1)$ ein r.R.-Körper über $\mathfrak{K}_0$, und das gleiche gilt für $\mathfrak{K}_0(\alpha_1, \alpha_2)$; denn $\mathfrak{K}_0(\gamma_1, \beta_1)$ bzw. $\mathfrak{K}_0(\alpha_1, \alpha_2)$ ist reell quadratisch über $\mathfrak{K}_0(\gamma_1)$ bzw. $\mathfrak{K}_0(\gamma_1, \beta_1)$. — Mit $\mathfrak{K}_0(\alpha_1, \alpha_2)$ ist auch $\mathfrak{K}_0(\alpha_1)$ ein r.R.-Körper über $\mathfrak{K}_0$. Aber man hat $[\mathfrak{K}_0(\alpha_1) : \mathfrak{K}_0] = 4$ und es läßt sich zwischen $\mathfrak{K}_0$ und $\mathfrak{K}_0(\alpha_1)$ kein echter Zwischenkörper einschieben.

Beachtet man, daß ein analoger Satz zu Satz 2 für jeden über $\mathfrak{K}_0$ normalen metazyklischen Körper gilt, während er für die Unterkörper von metazyklischen Normalkörpern nicht immer richtig ist, so erkennt man: *Die Maximalität ist bei den r.R.-Körpern ein gewisser Ersatz für die Normalität.* (Auf normale r.R.-Körper können wir uns nicht beschränken. Denn Hilfssatz 2 von § 2 zeigt, daß ein r.R.-Körper $\mathfrak{K}$ nur dann über $\mathfrak{K}_0$ normal sein kann, wenn $[\mathfrak{K} : \mathfrak{K}_0] = 2^m$.)

§2. Reelle Radikalkörper von Primzahlgrad

Hilfssatz 2. Ist $\mathfrak{K}$ r.R.-Körper über $\mathfrak{K}_0$ und $[\mathfrak{K} : \mathfrak{K}_0] = p \neq 2$ eine Primzahl, so sind alle echt Konjugierten von $\mathfrak{K}$ über $\mathfrak{K}_0$ komplex.

Hilfssatz 3. Ein reeller Abelscher Oberkörper des reellen Körpers $\mathfrak{K}_0$ von ungeradem Grad ist niemals r.R.-Körper über $\mathfrak{K}_0$.

Hilfssatz 2, der p-Konjugiertensatz der Einleitung, ist wohlbekannt. Hilfssatz 3 ist eine triviale Folgerung aus Hilfssatz 2. (Man beachte: Jeder Abelsche Oberkörper von $\mathfrak{K}_0$ enthält einen zyklischen Oberkörper mit $[\mathfrak{K} : \mathfrak{K}_0] = p$.) — Es sei jetzt $\mathfrak{K} = \mathfrak{K}_0(\alpha)$ ein fester Körper von ungeradem Primzahlgrad p über dem festen reellen Ausgangskörper $\mathfrak{K}_0$; $p(\alpha) = (x-\alpha_0) \dots (x-\alpha_{p-1})$. Sei das über $\mathfrak{K}_0$ irreduzible Polynom mit der reellen Nullstelle $\alpha = \alpha_0$, und es seien $\alpha_1, \dots, \alpha_{p-1}$ komplex. $\mathfrak{N} = \mathfrak{K}_0(\alpha_0, \dots, \alpha_{p-1})$ sei der zu $\mathfrak{K}$ über $\mathfrak{K}_0$ gehörige metazyklisch angenommene Normalkörper. Dann ist $\mathfrak{K}$ sicher nicht zyklisch über $\mathfrak{K}_0$. Nach der bekannten Theorie der metazyklischen Gleichungen von Primzahlgrad enthält $\mathfrak{N}$ einen über $\mathfrak{K}_0$ zyklischen Unterkörper $\mathfrak{Z}$, bei dem $t = [\mathfrak{Z} : \mathfrak{K}_0]$ ein Teiler von $p-1$, und für den $\mathfrak{N} = \mathfrak{Z} \cdot \mathfrak{K}_0(\alpha) = \mathfrak{Z} \cdot \mathfrak{K}_0(\alpha_i)$ $(i = 0, \dots, p-1;\ [\mathfrak{N} : \mathfrak{Z}] = p)$ gilt. Stellen wir die nicht-Abelsche Galoisgruppe $G(\mathfrak{N} : \mathfrak{K}_0)$ als Permutationsgruppe in den passend numerierten α_i dar, und setzen wir wie üblich $\alpha_k = \alpha_{k'}$, wenn $k \equiv k' (p)$, so besitzt $G(\mathfrak{N} : \mathfrak{K}_0)$ wie bekannt als Erzeugende die beiden Automorphismen $Z \equiv \{\alpha_i \to \alpha_{i+1}\}$ und $A \equiv \{\alpha_i \to \alpha_{s \cdot i}\}$; dabei ist s die kleinste positive Zahl mit den Eigenschaften $s^{t-1} \not\equiv 1 (p)$ und $s^t \equiv 1 (p)$. Der Invariantenkörper von Z bzw. A ist $\mathfrak{Z}$ bzw. $\mathfrak{K}_0(\alpha_0) = \mathfrak{K}$. Der durch A induzierte Automorphismus A_0 von $\mathfrak{Z}$ erzeugt die Gruppe $G(\mathfrak{Z} : \mathfrak{K}_0)$. Ist $[\mathfrak{Z} : \mathfrak{K}_0] = t = 2^m \cdot t_1$, $t_1 \equiv 2 (p)$, so besitzt $\mathfrak{Z}$ über $\mathfrak{K}_0$ eine direkte Zerlegung

$\mathfrak{Z} = \mathfrak{Z}_g \cdot \mathfrak{Z}_u$; $[\mathfrak{Z}_g : \mathfrak{K}_0] = 2^m$, $[\mathfrak{Z}_u : \mathfrak{K}_0] = t_1$. In entsprechender Weise läßt sich der über $\mathfrak{K}_0$ zyklische Körper $\mathfrak{K}_0(\varepsilon) = \mathfrak{E}\left(\varepsilon = \exp\frac{2\pi i}{p}\right)$ direkt zerlegen: $\mathfrak{E} = \mathfrak{E}_g \cdot \mathfrak{E}_u$; $[\mathfrak{E}_g : \mathfrak{K}_0] = 2^{m'}$, $[\mathfrak{E}_u : \mathfrak{K}_0] = t_1'$ $(t_1' \equiv 1(p),\ p - 1 \equiv 0(t_1'))$. Dabei sind $\mathfrak{Z}_u$ und $\mathfrak{E}_u$ reell; die Möglichkeiten $\mathfrak{Z}_u = \mathfrak{K}_0$, $\mathfrak{E}_u = \mathfrak{K}_0$ sind nicht ausgeschlossen, während sicher $\mathfrak{E}_g$ und $\mathfrak{Z}_g$ komplex, also von $\mathfrak{K}_0$ verschieden sind.

Satz 3. Ist $\mathfrak{K} = \mathfrak{K}_0(\alpha)$ ein r.R.-Körper über $\mathfrak{K}_0$, so gibt es einen reellen Quadratwurzelkörper $\mathfrak{K}_0'$ über $\mathfrak{K}_0$[8]*, derart daß $\mathfrak{N}' = \mathfrak{N} \cdot \mathfrak{K}_0' = \mathfrak{K}_0'(\varepsilon, \alpha)$; $[\mathfrak{N}' : \mathfrak{K}_0'] = 2pt_1$ $(t_1 \equiv 1(p))$.*

Den Beweis zerlegen wir in zwei Hauptabschnitte: a) *Es ist* $\mathfrak{Z}_u = \mathfrak{E}_u$. Es sei $\mathfrak{K}^*$ ein eigentlicher r.R.-Körper über $\mathfrak{K}_0$, der $\mathfrak{K}_0(\alpha)$ enthält; $\beta_1, \ldots, \beta_r$ sei eine erzeugende Radikalkette von $\mathfrak{K}^*$ und es sei $p(x)$ über $\mathfrak{L} = \mathfrak{K}(\beta_1, \ldots, \beta_{j-1})$ irreduzibel, über $\mathfrak{L}(\beta_j)$ reduzibel. Dann muß nach einer bekannten Verallgemeinerung von Hilfssatz 1 eine Nullstelle von $p(x)$ in $\mathfrak{L}(\beta_j)$ liegen[9], und das kann aus Realitätsgründen nur α sein. Weiter ergibt sich sofort: $\mathfrak{L}(\beta_j) = \mathfrak{L}(\alpha)$, $[\mathfrak{L}(\alpha) : \mathfrak{L}] = p$. Der zu $\mathfrak{L}(\alpha) = \mathfrak{L}(\beta_j)$ über $\mathfrak{L}$ gehörige Normalkörper $\mathfrak{N}_\mathfrak{L}$ ist einerseits gleich $\mathfrak{N} \cdot \mathfrak{L}$, andererseits gleich $\mathfrak{L}(\varepsilon)$. Daraus folgt $2^m \cdot t_1 \cdot p \equiv 0\ ([\mathfrak{N} \cdot \mathfrak{L} : \mathfrak{L}])$, $2^{m'} \cdot t_1' \cdot p \equiv 0([\mathfrak{N} \cdot \mathfrak{L} : \mathfrak{L}])$. Für den über $\mathfrak{K}_0$ Abelschen, reellen Körper $\mathfrak{Z}_u \cdot \mathfrak{E}_u$ haben wir $t_1 \cdot t_1' \equiv 0\ ([\mathfrak{Z}_u \cdot \mathfrak{E}_u : \mathfrak{K}_0])$. Nach Hilfssatz 3 ergibt sich daraus $\mathfrak{L} \cap \mathfrak{Z}_u \cdot \mathfrak{E}_u = \mathfrak{K}_0$ und damit $[\mathfrak{N} \cdot \mathfrak{L} : \mathfrak{K}_0] \equiv 0([\mathfrak{Z}_u \cdot \mathfrak{E}_u : \mathfrak{K}_0])$. Da nun $[\mathfrak{Z}_u \cdot \mathfrak{E}_u : \mathfrak{K}_0]$ durch das kleinste gemeinschaftliche Vielfache von t_1, t_1' teilbar und $t_1 \cdot t_1'$ zu $2p$ teilerfremd ist, sind die für $[\mathfrak{N} \cdot \mathfrak{L} : \mathfrak{L}]$ gefundenen Bedingungen nur dann miteinander verträglich, wenn $[\mathfrak{Z}_u : \mathfrak{K}_0] = [\mathfrak{E}_u : \mathfrak{K}_0] = [\mathfrak{Z}_u \cdot \mathfrak{E}_u : \mathfrak{K}_0]$ d.h. $\mathfrak{Z}_u = \mathfrak{E}_u$.

b) Es sei $\mathfrak{K}_0'$ größter reeller Unterkörper von $\mathfrak{Z}_g \cdot \mathfrak{E}_g$, Dann ist $[\mathfrak{Z}_g \cdot \mathfrak{E}_g : \mathfrak{K}_0'] = 2$ wegen der Normalität von $\mathfrak{K}_0'$ über dem reellen Körper $\mathfrak{K}_0$, und da $\mathfrak{K}_0'$ über $\mathfrak{K}_0$ Abelsch und reell, existiert für $\mathfrak{K}_0'$ über $\mathfrak{K}_0$ eine erzeugende Radikalkette $\varrho_1, \ldots, \varrho_m$, bei der $[\mathfrak{K}_0(\varrho_1, \ldots, \varrho_i) : \mathfrak{K}_0(\varrho_1, \ldots, \varrho_{i-1})] = 2$ $(i = 1, \ldots, m)$. Von den komplexen Körpern $\mathfrak{Z}_g$, $\mathfrak{E}_g$ ist keiner in $\mathfrak{K}_0'$ enthalten; es ist also $\mathfrak{Z}_g \cdot \mathfrak{K}_0' = \mathfrak{E}_g \cdot \mathfrak{K}_0' = \mathfrak{Z}_g \cdot \mathfrak{E}_g$. — Wir haben jetzt unter Berücksichtigung von $\mathfrak{E}_u = \mathfrak{Z}_u$: $\varepsilon \in \mathfrak{K}_0'(\alpha_0) \cdot \mathfrak{Z} = \mathfrak{K}_0' \cdot \mathfrak{N} = \mathfrak{N}'$, $\mathfrak{K}_0'(\varepsilon, \alpha) \subseteq \mathfrak{N}'$; $\mathfrak{Z} \subseteq \mathfrak{K}_0'(\varepsilon) = \mathfrak{K}_0' \cdot \mathfrak{E}$, $\mathfrak{N}' \subseteq \mathfrak{K}_0'(\varepsilon, \alpha)$, $\mathfrak{N}' = \mathfrak{K}_0(\varepsilon, \alpha)$. Schließlich ist $[\mathfrak{N}' : \mathfrak{K}_0'] = [\mathfrak{N}' : \mathfrak{K}_0'(\varepsilon)] \cdot [\mathfrak{K}_0'(\varepsilon) : \mathfrak{K}_0'] = p \cdot [\mathfrak{K}_0'(\varepsilon) : \mathfrak{K}_0']$; $[\mathfrak{K}_0'(\varepsilon) : \mathfrak{K}_0'] = [\mathfrak{K}_0'(\varepsilon) : \mathfrak{E}_u \cdot \mathfrak{K}_0'] \cdot [\mathfrak{E}_u \cdot \mathfrak{K}_0' : \mathfrak{K}_0'] = [\mathfrak{K}_0'(\varepsilon) : \mathfrak{E}_u \cdot \mathfrak{K}_0'] \cdot t_1$; $[\mathfrak{K}_0'(\varepsilon) : \mathfrak{E}_u \cdot \mathfrak{K}_0'] = [\mathfrak{K}_0' \cdot \mathfrak{E}_u \cdot \mathfrak{E}_g : \mathfrak{E}_u \cdot \mathfrak{K}_0'] = [\mathfrak{K}_0' \cdot \mathfrak{E}_g : \mathfrak{K}_0'] = [\mathfrak{Z}_g \cdot \mathfrak{E}_g : \mathfrak{K}_0'] = 2$, also $[\mathfrak{N}' : \mathfrak{K}_0'] = 2p\,t_1$.

[8]) Also ein reeller Oberkörper $\mathfrak{K}$ von $\mathfrak{K}_0$, zu dem eine Kette $\mathfrak{K}_0 < \mathfrak{K}_1 < \cdots < \mathfrak{K}_m = \mathfrak{K}$ mit $[\mathfrak{K}_i : \mathfrak{K}_{i-1}] = 2$ $(i = 1, \ldots, m)$ existiert, so daß jeweils $\mathfrak{K}_i = \mathfrak{K}_{i-1}(\sqrt{a_i})$ $(a_i \in \mathfrak{K}_{i-1},\ a_i > 0)$. Man beachte: a) Ein reeller Quadratwurzelkörper $\mathfrak{K}$ über $\mathfrak{K}_0$ ist stets ein r.Q.-Körper über $\mathfrak{K}_0$ mit $[\mathfrak{K} : \mathfrak{K}_0] = 2^m$; aber ein r.R.-Körper $\mathfrak{K}$ mit $[\mathfrak{K} : \mathfrak{K}_0] = 2^m$ braucht nicht immer ein reeller Quadratwurzelkörper über $\mathfrak{K}_0$ zu sein (vgl. das im Anschluß an Satz 2 von § 1 angegebene Beispiel). b) Ein maximaler r.R.-Körper $\mathfrak{K}$ mit $[\mathfrak{K} : \mathfrak{K}_0] = 2^m$ ist nach Satz 2 stets ein reeller Quadratwurzelkörper. c) Ein reeller Quadratwurzelkörper über $\mathfrak{K}_0$ kann auch charakterisiert werden, als ein reeller Unterkörper eines Normaloberkörpers $\mathfrak{N}$ von $\mathfrak{K}_0$, bei dem $[\mathfrak{N} : \mathfrak{K}_0] = 2^n$. (Unschwer einzusehen, wenn man beachtet, daß bei einem $\mathfrak{N}$ mit $[\mathfrak{N} : \mathfrak{K}_0] = 2^n$ wegen der Hyperauflösbarkeit der Gruppe zu jedem Unterkörper $\mathfrak{K}$ eine Kette $\mathfrak{K}_0 < \mathfrak{K}_1 < \cdots < \mathfrak{K}_m = \mathfrak{K}$ mit $[\mathfrak{K}_i : \mathfrak{K}_{i-1}] = 2$ existiert.)

[9]) Folgt aus der Tatsache, daß jede intransitive Untergruppe von $G(\mathfrak{N} : \mathfrak{K}_0)$ mindestens ein α_i fest lassen muß.

$\mathfrak{N}'$ baut sich über $\mathfrak{K}_0'$ nach dem gleichen Schema auf wie $\mathfrak{N}$ über $\mathfrak{K}_0$. Der wesentliche Unterschied ist nur der, daß an Stelle von $\mathfrak{Z}$ der Körper $\mathfrak{Z}' = \mathfrak{K}_0'(\varepsilon)$ tritt, und daß $[\mathfrak{Z}' : \mathfrak{K}_0'] = 2t_1$, statt $[\mathfrak{Z} : \mathfrak{K}_0] = 2^m \cdot t_1$ gilt. Wir betrachten nun unter Weglassung der $'$ und unter vorläufigem Verzicht auf die Gradbedingung allgemein den Fall, daß $\mathfrak{N} = \mathfrak{K}_0(\alpha) \cdot \mathfrak{Z} = \mathfrak{K}_0(\varepsilon, \alpha)$, $\mathfrak{Z} = \mathfrak{K}_0(\varepsilon)$ wird. Um hier eine notwendige und hinreichende Bedingung dafür zu erhalten, daß $\mathfrak{K}_0(\alpha)$ einen r. R.-Körper über $\mathfrak{K}_0$ darstellt, führen wir eine Zahlinvariante k von $\mathfrak{N}$ über $\mathfrak{K}_0$ ein, die aus historischen Gründen als die *Kummersche Invariante* bezeichnet werden kann[10]). Da $\mathfrak{N}$ über $\mathfrak{Z}$ zyklisch ist und ε in $\mathfrak{Z}$ liegt, gibt es für $\mathfrak{N}$ über $\mathfrak{Z}$ ein erzeugendes Binom $x^p - b$ ($b \in \mathfrak{Z}$), es wird also $\mathfrak{N} = \mathfrak{Z}(\beta)$, $\beta^p = b$. Bilden wir β^A mit dem oben eingeführten Automorphismus A, so muß wegen der Normalität von $\mathfrak{N}$ über $\mathfrak{K}_0$ auch $\mathfrak{N} = \mathfrak{Z}(\beta^A)$ werden, und daraus folgt eine leichte Rechnung $\beta^A = c \cdot \beta^k$ ($1 \leq k \leq p-1$) in $\mathfrak{N}$, also $b^{A^\varrho} = c^p \cdot \beta^{k^\varrho}$ ($c \in \mathfrak{Z}$) in $\mathfrak{Z}$. Der Exponent k, von dem man mühelos zeigt, daß er nicht von der speziellen Wahl von b, sondern nur von $\mathfrak{N}$ und $\mathfrak{K}_0$ abhängt, ist die Kummersche Invariante von $\mathfrak{N} = \mathfrak{K}_0(\alpha, \varepsilon)$ über $\mathfrak{K}_0$. Aus der Definition von k folgt sofort:

Ist $\mathfrak{K}_0' \supset \mathfrak{K}_0$ und $\mathfrak{K}_0' \cap \mathfrak{N} = \mathfrak{K}_0$, so hat $\mathfrak{N}' = \mathfrak{K}_0' \cdot \mathfrak{N} = \mathfrak{K}_0'(\alpha, \varepsilon)$ über $\mathfrak{K}_0'$ dieselbe Kummersche Invariante k wie $\mathfrak{N}$ über $\mathfrak{K}_0$.

Man beachte: Wegen $\mathfrak{N} \cap \mathfrak{K}_0' = \mathfrak{K}_0$ sind die Gruppen $G(\mathfrak{N} : \mathfrak{K}_0)$ und $G(\mathfrak{N}' : \mathfrak{K}_0')$ isomorph und es ist die A zugeordnete Fortsetzung A' von A auf $\mathfrak{N}'$ für die Berechnung der Kummerschen Invariante k bei $\mathfrak{N}'$, $\mathfrak{K}_0'$ genau so brauchbar wie A bei $\mathfrak{N}$, $\mathfrak{K}_0$. Ist ferner $\mathfrak{N} = \mathfrak{Z}(\beta)$, $\beta^p = b \in \mathfrak{Z}$, so ist $\mathfrak{N}' = \mathfrak{Z}'(\beta')$ ($\mathfrak{Z}' = \mathfrak{Z} \cdot \mathfrak{K}_0'$) und aus $\beta = c \cdot \beta^A = c \cdot \beta^{A'}$, $c \in \mathfrak{Z} \subseteq \mathfrak{Z}'$ folgt sofort die Behauptung.

Satz 4. a) Hat $\mathfrak{N} = \mathfrak{K}_0(\alpha, \varepsilon)$ über $\mathfrak{K}_0$ die Kummersche Invariante $k = 1$, so ist $\mathfrak{K}_0(\alpha)$ eigentlicher r. R.-Körper über $\mathfrak{K}_0$. b) Ist $\mathfrak{N} = \mathfrak{K}_0(\alpha, \varepsilon)$, $[\mathfrak{N} : \mathfrak{K}_0] = 2p\,t_1$ $(t_1 \equiv 1(2))$ und ist $\mathfrak{K}_0(\alpha)$ r. R.-Körper über $\mathfrak{K}_0$, so hat $\mathfrak{N}$ über $\mathfrak{K}_0$ die Kummersche Invariante 1.

Wir beginnen mit b). Ist $\mathfrak{K}_0(\alpha)$ r. R.-Körper über $\mathfrak{K}_0$, so gibt es nach einer schon beim Beweise von Satz 3 benutzten Überlegung einen r. R.-Körper $\mathfrak{K}_0'$ über $\mathfrak{K}_0$, derart, daß $[\mathfrak{K}_0'(\alpha) : \mathfrak{K}_0'] = p$ und gleichzeitig $\mathfrak{K}_0'(\alpha) = \mathfrak{K}_0'(\beta)$, wobei β Nullstelle eines über $\mathfrak{K}_0'$ irreduziblen Binoms $x^p - b$ ($b \in \mathfrak{K}_0'$). Aus $[\mathfrak{K}_0'(\alpha) : \mathfrak{K}_0] = p$, d. h. der Irreduzibilität von $p(x) = (x - \alpha_0) \cdot \cdots \cdot (x - \alpha_{p-1})$ über $\mathfrak{K}_0'$ folgt nach der bekannten Theorie der metazyklischen Gleichungen $\mathfrak{K}_0' \cap \mathfrak{N} = \mathfrak{K}_0' \cap \mathfrak{Z}$. Wegen $[\mathfrak{Z} : \mathfrak{K}_0] = [\mathfrak{Z}_g \cdot \mathfrak{Z}_u : \mathfrak{K}_0] = 2t_1$ muß aber $\mathfrak{K}_0' \cap \mathfrak{Z} = \mathfrak{K}_0$ sein; denn es ist ja $\mathfrak{K}_0' \cap \mathfrak{Z}_u = \mathfrak{K}_0$ nach Hilfssatz 3, und $\mathfrak{K}_0' \cap \mathfrak{Z}_g = \mathfrak{K}_0$, weil $\mathfrak{Z}_g$ ein komplexer Oberkörper zweiten Grades von $\mathfrak{K}_0$. Wegen $\mathfrak{K}_0' \cap \mathfrak{N} = \mathfrak{K}_0$ sind $G(\mathfrak{N} : \mathfrak{K}_0)$ und $G(\mathfrak{N} \cdot \mathfrak{K}_0' : \mathfrak{K}_0')$ isomorph und die Kummerschen Invarianten von $\mathfrak{N}$ über $\mathfrak{K}$ und $\mathfrak{N}'$ über $\mathfrak{K}_0'$ gleich. Die Kummersche Invariante von $\mathfrak{N}'$ über $\mathfrak{K}_0'$ schließlich ist offenbar gleich 1, da $b^{A'} = b$, $\beta^{A'} = \varepsilon^\varrho \cdot \beta$, $\varepsilon^\varrho \in \mathfrak{Z} \subseteq \mathfrak{Z} \cdot \mathfrak{K}_0'$.

Zum Beweise von a) genügt es, zu zeigen, daß ein reelles $\lambda \in \mathfrak{N}$ existiert, derart daß $\lambda^p = b \in \mathfrak{K}_0$ und $x^p - b$ irreduzibel über $\mathfrak{K}_0$. Denn dann muß

[10]) Zur Kummerschen Invariante vgl. etwa HILBERT, *Zahlbericht* (Gesammelte Abhandlungen I, 7.) § 125, Satz 147.

$p(x) = (x-\alpha_0)\cdot\cdots\cdot(x-\alpha_{p-1})$ über $\mathfrak{K}_0(\lambda)$ reduzibel werden; nach einer bereits bei Satz 3 benutzten Verallgemeinerung von Hilfssatz 1 muß also $\mathfrak{K}_0(\lambda)$ eine Nullstelle von $p(x)$ enthalten, und wegen der Realität von $\mathfrak{K}_0(\lambda)$ kann das nur $\alpha = \alpha_0$ sein[11]). — Wir bilden nun bei der zu Beginn des Paragraphen festgelegten Anordnung der α_i die Lagrangeschen Resolventen

$$(1) \qquad \lambda_q = \alpha_0 + \sum_{k=1}^{p-1} \varepsilon^{q\cdot k}\cdot\alpha_k \qquad (\lambda_q = \lambda_{q'},\ \text{für}\ \ q \equiv q'(p)).$$

Ohne Beschränkung der Allgemeinheit dürfen wir $\lambda_1 \neq 0$ annehmen[12]). Wegen $\mathfrak{Z} = \mathfrak{K}_0(\varepsilon)$ ist $\varepsilon^Z = \varepsilon$, $\varepsilon^{A_0} = \varepsilon^A = \varepsilon^{s'}$, wobei $s'^{t-1} = s'^{2t_1-1} \not\equiv 1(p)$, $s'^t = s'^{2t_1} \equiv 1(p)$, weil A_0 erzeugender Automorphismus von $G(\mathfrak{Z}:\mathfrak{K}_0)$. Mit $r\cdot s \equiv 1(p)$ ergibt sich:

$$(2) \qquad \lambda_1^A = \alpha_0 + \sum_{k=1}^{p-1} \varepsilon^{s'\cdot k}\cdot\alpha_{s\cdot k} = \alpha_0 + \sum_{k=1}^{p-1} \varepsilon^{s'\cdot r\cdot k}\cdot\alpha_k = \lambda_{s'\cdot r}.$$

(Man darf ja $s\cdot k$ durch k ersetzen, wenn man gleichzeitig $r\cdot k$ an Stelle von k schreibt.) Wegen $\lambda_q^Z = \varepsilon^{-q}$ und (2) haben wir: $(\lambda_1^A\cdot\lambda_1^{-k})^Z = \lambda_1^A\cdot\lambda_1^{-k}$, $\lambda_1^A\cdot\lambda_1^k \in \mathfrak{Z}$ für $k \equiv s'\cdot r(p)$, d.h.: Die Kummersche Invariante k von $\mathfrak{K}$ über $\mathfrak{K}_0$ ist charakterisiert durch $1 \leq k \leq p-1$, $k \equiv s'\cdot r(p)$. Da nun $k=1$ nach Voraussetzung, muß $s'\cdot r \equiv 1(p)$, $\lambda_1^A = \lambda_1$ sein. Für $\lambda_1^p = b$ ergibt sich weiter: $b^A = b$, $b^Z = b$; $b \in \mathfrak{K}_0$. Weil schließlich $x^p - \lambda_1^p = x^p - b$ irreduzibel über $\mathfrak{K}_0$ ist und $\mathfrak{N}$ als Normalkörper über $\mathfrak{K}_0$ gleichzeitig mit λ_1 auch die reelle Nullstelle λ von $x^p - b$ enthält, ist nach einer Bemerkung von weiter oben mit der Feststellung $b \in \mathfrak{K}_0$ der Beweis von Satz 4a) abgeschlossen. — Aus Satz 3 und Satz 4 zusammen ergibt sich unmittelbar:

I. Hauptsatz der r.R.-Körper von Primzahlgrad. Ein Oberkörper $\mathfrak{K}_0(\alpha)$ des reellen Körpers $\mathfrak{K}_0$ von ungeradem Primzahlgrad p ist dann und nur dann ein r.R.-Körper über $\mathfrak{K}_0$, wenn sich ein reeller Quadratwurzelkörper $\mathfrak{K}_0'$ über $\mathfrak{K}_0$ so bestimmen läßt, daß für den zu $\mathfrak{K}_0'(\alpha)$ über $\mathfrak{K}_0'$ gehörigen Normalkörper $\mathfrak{N}'$ die Gleichungen $\mathfrak{N}' = \mathfrak{K}_0'(\alpha, \varepsilon)$ $\left(\varepsilon = \exp\frac{2\pi i}{p}\right)$, $[\mathfrak{N}':\mathfrak{K}_0'] = 2p\,t_1$ $(t_1 \equiv 1(2))$ gelten, und wenn dabei außerdem $\mathfrak{N}'$ über $\mathfrak{K}_0'$ die Kummersche Invariante $k=1$ besitzt.

Sind nämlich die Voraussetzungen des Hauptsatzes erfüllt, so ist $\mathfrak{K}_0'(\alpha')$ nach Satz 4 r.R.-Körper über $\mathfrak{K}_0'$ und damit auch über $\mathfrak{K}_0$, da ja $\mathfrak{K}_0'$ seinerseits r.R.-Körper über $\mathfrak{K}_0$. Als Unterkörper von $\mathfrak{K}_0'(\alpha)$ ist also $\mathfrak{K}_0(\alpha)$ r.R.-Körper über $\mathfrak{K}_0$. — Ist andererseits $\mathfrak{K}_0(\alpha)$ r.R.-Körper über $\mathfrak{K}_0$, so existiert nach Satz 3 ein Quadratwurzelkörper $\mathfrak{K}_0'$ der im Hauptsatz geforderten Art und es ist $\mathfrak{K}_0'(\alpha)$ nach § 1 r.R.-Körper über $\mathfrak{K}_0'$. Nach Satz 4b) muß also $\mathfrak{N}'$ über $\mathfrak{K}_0'$ die Kummersche Invariante $k=1$ haben. — Am interessantesten ist beim ersten Hauptsatz die Feststellung, daß $\mathfrak{K}_0(\alpha)$ auch im Falle $\mathfrak{N} = \mathfrak{K}_0(\alpha, \varepsilon)$

[11]) Daß außer α_0 keine weitere Nullstelle reell sein kann, folgt ohne Zusatzvoraussetzung aus der Tatsache, daß $\mathfrak{N}$ komplex ist. Denn nach einem bekannten Satz wird $\mathfrak{N} = \mathfrak{K}_0(\alpha_0, \alpha_i)$, falls α_i beliebig aus der Reihe $\alpha_i, \ldots, \alpha_{p-1}$ gewählt.

[12]) Wir müssen dann allerdings unter Umständen für ε statt $\exp\left(\frac{2\pi i}{p}\right)$ eine andere primitive p-te Einheitswurzel wählen; doch ist das völlig nebensächlich.

keineswegs immer r.R.-Körper über $\mathfrak{K}_0$ zu sein braucht. (Man kann mit zahlentheoretischen Überlegungen leicht Beispiele konstruieren, in denen $\mathfrak{N}=\mathfrak{K}_0(\alpha,\varepsilon)$, $[\mathfrak{N}:\mathfrak{K}_0]=2pt_1$, aber $k\neq 1$.) Einen wichtigen Sonderfall gibt es, in dem sich die Bedingungen von Hauptsatz 1 außerordentlich vereinfachen lassen. Es sei p eine *Gaußsche Primzahl*, also $p=2^N+1$. Dann wird $t_1=t$, $\mathfrak{E}_u=\mathfrak{Z}_0=\mathfrak{K}_0$, und es ist die Existenz des Quadratwurzelkörpers $\mathfrak{K}_0'$ nach dem Beweis von Satz 3 gesichert, sofern nur $\mathfrak{N}=\mathfrak{K}_0(\alpha_0,\alpha_1,\ldots,\alpha_{p-1})$ komplex. Weiter muß aber bei einer Gaußschen Primzahl die Kummersche Invariante k von $\mathfrak{N}'$ über $\mathfrak{K}_0'$ stets gleich 1 sein. Denn wir haben hier $s\not\equiv 1(p)$, $s'\not\equiv 1(p)$, $s^2\equiv s'^2\equiv 1(p)$, also $s\equiv s'\equiv -1(p)$, und aus $s\cdot r\equiv 1(p)$ folgt $s'\cdot r\equiv k\equiv 1(p)$. Für Gaußsche Primzahlen gilt also:

II. Hauptsatz der r.R.-Körper von Primzahlgrad. Ist $\mathfrak{K}_0(\alpha)$ reeller metazyklischer Oberkörper von $\mathfrak{K}_0$ und ist $[\mathfrak{K}_0(\alpha):\mathfrak{K}_0]$ eine Gaußsche Primzahl p, so ist $\mathfrak{K}_0(\alpha)$ dann und nur dann ein r.R.-Körper über $\mathfrak{K}_0$, wenn der zu $\mathfrak{K}_0(\alpha)$ über $\mathfrak{K}_0$ gehörige Normalkörper $\mathfrak{N}$ komplex ist.

(Zu unserer Formulierung der notwendigen Bedingung — im Vergleich zu Hilfssatz 3 — beachte man: Ist $\mathfrak{N}$ komplex, so müssen nach einem bekannten, auch weiter oben gelegentlich benutzten Satz[11]) *alle* echt Konjugierten von α komplex sein.) Hauptsatz II zeigt, daß 7 die kleinste Primzahl ist, bei der die eigentliche Problematik der r.R.-Körper von Primzahlgrad erkennbar wird.

§3. Beliebige maximale r.R.-Körper. Der Konjugiertensatz

Ein reeller Quadratwurzelkörper über $\mathfrak{K}_0$[8]) möge kurz als r.Q.-Körper über $\mathfrak{K}_0$ bezeichnet werden. Man beachte: Ist $\mathfrak{K}$ r.Q.-Körper über $\mathfrak{K}_0$, $\mathfrak{M}$ ein beliebiger reeller Oberkörper von $\mathfrak{K}_0$, so ist $\mathfrak{K}$ auch r.Q.-Körper über $\mathfrak{M}$.

Satz 5. Ist $\mathfrak{K}$ ein maximaler r.R.-Körper über $\mathfrak{K}_0$ mit dem zugehörigen Normalkörper $\mathfrak{N}$, und sind $p_1,\ldots,p_l$ die in $[\mathfrak{K}:\mathfrak{K}_0]$ aufgehenden ungeraden Primzahlen, so gibt es einen r.Q.-Körper $\mathfrak{K}^*$ über $\mathfrak{K}$, derart daß $\varepsilon_k\in\mathfrak{N}\cdot\mathfrak{K}^*$ $\left(\varepsilon_k=\exp\frac{2\pi i}{k};\ k=1,\ldots,l\right)$.

Nach Satz 2 existieren für jede Primzahl p_k zwei Körper $\mathfrak{K}_{1k}$, $\mathfrak{K}_{2k}$, derart daß $\mathfrak{K}_0\subseteqq\mathfrak{K}_{1k}\subset\mathfrak{K}_{2k}\subseteqq\mathfrak{K}$, $[\mathfrak{K}_{2k}:\mathfrak{K}_{1k}]=p_k$. Da dabei $\mathfrak{K}_{2k}$ r.R.-Körper über $\mathfrak{K}_{1k}$, gibt es nach Satz 3 einen r.Q.-Körper $\mathfrak{K}_k^*$ über $\mathfrak{K}_{1k}$, derart daß $\varepsilon_k\in\mathfrak{N}_k\cdot\mathfrak{K}_k^*$, falls $\mathfrak{N}_k$ der zu $\mathfrak{K}_{2k}$ über $\mathfrak{K}_{1k}$ gehörige, in $\mathfrak{N}$ enthaltene Normalkörper. Setzt man nun $\mathfrak{K}^*=\mathfrak{K}_1^*\cdot\cdots\cdot\mathfrak{K}_l^*$, so ist $\mathfrak{K}^*$ ein r.Q.-Körper über $\mathfrak{N}$ und $\varepsilon_k\in\mathfrak{N}\cdot\mathfrak{K}^*$ $(k=1,\ldots,l)$.

Satz 6. Zu jedem maximalen r.R.-Körper $\mathfrak{K}$ über $\mathfrak{K}_0$ gibt es einen r.Q.-Körper $\mathfrak{K}^*$ über $\mathfrak{K}$, derart daß $\mathfrak{K}^*$ eigentlicher r.R.-Körper über $\mathfrak{K}_0$.

Für $[\mathfrak{K}:\mathfrak{K}_0]=2$ ist die Behauptung trivial, für $[\mathfrak{K}:\mathfrak{K}_0]=p\neq 2$ (p Primzahl) folgt sie aus Satz 2 und Satz 3a) von § 2. [Man beachte: Ist $\mathfrak{K}_0'$ ein r.Q.-Körper über $\mathfrak{K}_0$, $\mathfrak{K}_0'(\alpha)$ eigentlicher r.R.-Körper über $\mathfrak{K}_0'$, so ist $\mathfrak{K}_0'(\alpha)$ eigentlicher r.R.-Körper über $\mathfrak{K}_0$ und r.Q.-Körper über $\mathfrak{K}_0'$.] — Im allgemeinen Fall gibt es nach Satz 2 jedenfalls einen Körper $\mathfrak{K}_1$, derart daß $[\mathfrak{K}_1:\mathfrak{K}_0]$ Primzahl (gerade oder ungerade). Wie bereits besprochen, existiert ein $\mathfrak{K}_1^*$, der r.Q.-Körper über $\mathfrak{K}_1$ und eigentlicher r.R.-Körper über $\mathfrak{K}_0$. Da ferner $\mathfrak{K}$ ein

maximaler r.R.-Körper über $\mathfrak{K}_1$ und da $[\mathfrak{K}:\mathfrak{K}_1]$ weniger Primfaktoren enthält als $[\mathfrak{K}:\mathfrak{K}_0]$, dürfen wir nach Induktionsprinzipien annehmen, daß es einen r.Q.-Körper $\mathfrak{K}_2^*$ über $\mathfrak{K}$ gibt, der eigentlicher r.R.-Körper über $\mathfrak{K}_1$ ist. $\mathfrak{K}_1^* \cdot \mathfrak{K}_2^* = \mathfrak{K}^*$ ist r.Q.-Körper über $\mathfrak{K}$, und die beim Beweise von Satz 1c) benutzten Überlegungen zeigen, daß $\mathfrak{K}^*$ ein eigentlicher r.R.-Körper über $\mathfrak{K}_0$ sein muß. — Daß bei Satz 5 und 6 die Beschränkung auf *maximale* r.R.-Körper wesentlich war, folgt aus der Bemerkung, daß das entscheidende Beweishilfsmittel der nur für maximale r.R.-Körper gültige Satz 2 war.

Satz 5 und Satz 6 legen den Gedanken nahe, es müßten sich wenigstens die maximalen r.R.-Körper ohne zusätzliche Einschränkung ebenso erschöpfend charakterisieren lassen wie die r.R.-Körper von Primzahlgrad. Indessen zeigt es sich rasch, daß dieses *Zentralproblem der r.R.-Körper* sicher keine einfache Aufgabe ist, und daß man jedenfalls nicht mit so primitiven Induktionsschlüssen zum Ziel kommt, wie bei Satz 5 und 6. Die Hauptschwierigkeit liegt darin, daß ein durch Übereinandertürmung relativer Normalkörper entstehender Körper nicht mehr über dem ursprünglichen Ausgangskörper normal zu sein braucht[13]). Wir müssen uns daher damit begnügen, wenigstens ein erstes, allgemeines Theorem zu beweisen:

Satz 7. Konjugiertensatz für maximale r.R.-Körper. Ist $\mathfrak{K}$ ein maximaler r.R.-Körper über $\mathfrak{K}_0$, so sind die zu $\mathfrak{K}$ über $\mathfrak{K}_0$ echt konjugierten Körper[2]) *alle komplex.*

Für $[\mathfrak{K}:\mathfrak{K}_0] = 2$ ist Satz 7 trivialerweise richtig; denn dann ist $\mathfrak{K}$ über $\mathfrak{K}_0$ normal und es gibt überhaupt keine echt konjugierten Körper zu $\mathfrak{K}$. Ist $[\mathfrak{K}:\mathfrak{K}_0]$ eine ungerade Primzahl p, so folgt die Richtigkeit von Satz 7 aus Hilfssatz 2. Im allgemeinen Fall sind wir am Ziel, wenn wir zeigen können: Gibt es einen echt konjugierten Körper zu $\mathfrak{K}$, so existiert stets ein über $\mathfrak{K}_0$ nicht normaler Körper $\mathfrak{K}_0'' \subseteq \mathfrak{K}$ mit der Eigenschaft, daß alle zu $\mathfrak{K}_0''$ über $\mathfrak{K}_0$ echt konjugierten Körper komplex sind. Denn dann ist, falls $\mathfrak{N}$ der zu $\mathfrak{K}$ über $\mathfrak{K}_0$ gehörige Normalkörper, der Normalisator von $G(\mathfrak{N}:\mathfrak{K}_0'')$ wegen der Nichtnormalität von $\mathfrak{K}_0''$ eine echte Untergruppe von $G(\mathfrak{N}:\mathfrak{K}_0)$, es ist also der Invariantenkörper $\mathfrak{K}_0'$ dieses Normalisators ein echter Oberkörper von $\mathfrak{K}_0$. Ferner ist klar, daß in unserem Falle alle zu $\mathfrak{K}$ über $\mathfrak{K}_0$ echt konjugierten Körper komplex sind, wenn nur die zu $\mathfrak{K}$ über $\mathfrak{K}_0'$ echt konjugierten Körper diese Eigenschaft haben. Das aber dürfen wir nach Induktionsprinzip als gesichert annehmen, da $\mathfrak{K}$ auch ein maximaler r.R.-Körper über $\mathfrak{K}_0'$ und $[\mathfrak{K}:\mathfrak{K}_0']$ ein echter Teiler von $[\mathfrak{K}:\mathfrak{K}_0]$. — Bei der Konstruktion von $\mathfrak{K}_0''$ unterscheiden wir mehrere Fälle:

a) Es sei $\mathfrak{K}$ ein Quadratwurzelkörper über $\mathfrak{K}_0$, also $[\mathfrak{K}:\mathfrak{K}_0] = 2^s$ und es sei $\mathfrak{N}_0$ der größte in $\mathfrak{K}$ enthaltene, über $\mathfrak{K}_0$ normale Körper. Für $\mathfrak{N}_0 = \mathfrak{K} = \mathfrak{N}$ ist Satz 7 trivial richtig. Sei also $\mathfrak{N}_0 \subset \mathfrak{K}$; da $\mathfrak{K}$ ein maximaler r.R.-Körper

[13]) Daß man möglicherweise diese Schwierigkeit dadurch umgehen kann, daß man von den maximalen r.R.-Körpern zu den zugehörigen Normalkörpern übergeht, wurde am Schluß der Einleitung angedeutet. Allerdings dürfte auch die Charakterisierung dieser Normalkörper keineswegs einfach sein. Am besten beginnt man vielleicht mit der Betrachtung von Spezialfällen, etwa mit dem Studium der komplizierteren Normalkörper, die unter e) bei dem Induktionsbeweis von Satz 7 auftreten.

über $\mathfrak{R}_0$, existiert ein Körper $\mathfrak{R}_1$ mit $\mathfrak{R}_0 < \mathfrak{R}_1 \leqq \mathfrak{R}$; $[\mathfrak{R}_1 : \mathfrak{R}_0] = 2$. Der zu $\mathfrak{R}_1$ über $\mathfrak{R}_0$ gehörige Normalkörper $\mathfrak{N}_1$ ist das Produkt von endlich vielen über $\mathfrak{N}_0$ quadratischen Körpern $\mathfrak{R}_1 = \mathfrak{R}', \mathfrak{R}'', \ldots$ und es ist $\mathfrak{N}_1$ sicher komplex, weil $\mathfrak{N}_1$ im Falle der Realität wegen der Maximaleigenschaft von $\mathfrak{R}$ in $\mathfrak{R}$ enthalten sein müßte, was der Definition von $\mathfrak{N}_0$ widerspricht. Es seien nun etwa $\mathfrak{R}_1 = \mathfrak{R}', \ldots, \mathfrak{R}^{(\varrho)}$ reell, $\mathfrak{R}^{(\varrho+1)}, \mathfrak{R}^{(\varrho+2)}, \ldots$ komplex, $\mathfrak{R}_0'' = \mathfrak{R}' \cdot \mathfrak{R}'' \cdot \ldots \cdot \mathfrak{R}^{(\varrho)}$. Da jeder Automorphismus A_1 von $\mathfrak{N}_1$ über $\mathfrak{R}_0$ einen Automorphismus A_0 von $\mathfrak{N}_0$ induziert, vertauschen sich bei jedem A_1 die Körper $\mathfrak{R}', \mathfrak{R}'', \ldots$ untereinander und man hat zwei Möglichkeiten: Entweder es geht bei A_1 jedes $\mathfrak{R}^{(\nu)}$ $(\nu \leq \varrho)$ wieder in ein $\mathfrak{R}^{(\mu)}$ $(\mu \leq \varrho)$ über; dann induziert A_1 einen Automorphismus von $\mathfrak{R}_0''$. Oder es hat für mindestens ein $\nu \leq \varrho$ das Bild von $\mathfrak{R}^{(\nu)}$ relativ A_1 einen Index $\mu > \varrho$; dann bilden A_1 den Körper $\mathfrak{R}_0''$ auf einen komplexen Körper ab. Die echt Konjugierten von $\mathfrak{R}_0''$ über $\mathfrak{R}_0$ sind also alle komplex. Außerdem ist $\mathfrak{R}_0''$ als Produkt von r. Q.-Körpern selbst ein r. Q.-Körper über $\mathfrak{N}_0$ und damit auch ein r. R.-Körper über $\mathfrak{R}_0$.

b) Offenbar gibt es in $\mathfrak{R}$ einen größten r. Q.-Körper $\mathfrak{R}_q$, den Vereinigungskörper aller der $\mathfrak{Q} \leqq \mathfrak{R}$, die über $\mathfrak{R}_0$ r. Q.-Körper sind. Ist $\mathfrak{R}_q = \mathfrak{R}_0$, so enthält $\mathfrak{R}$ einen Körper $\mathfrak{R}_1$, dessen Grad über $\mathfrak{R}_0$ eine ungerade Primzahl p ist. Hilfssatz 2 von § 2 zeigt, daß $\mathfrak{R}_1$ über $\mathfrak{R}_0$ nicht normal sein kann, und daß alle echt konjugierten Körper von $\mathfrak{R}_1$ komplex sind. Wir können also $\mathfrak{R}_0'' = \mathfrak{R}_1$ setzen und sind am Ziel.

c) Es sei $\mathfrak{R}_q > \mathfrak{R}_0$, aber von dem zugehörigen Normalkörper $\mathfrak{N}_q$ (über $\mathfrak{R}_0$) verschieden. Dann ist, wie sofort zu sehen, der Grad $[\mathfrak{N}_q : \mathfrak{R}_0]$ eine Potenz von 2. Der größte in $\mathfrak{N}_q$ enthaltene r. R.-Körper $\mathfrak{R}_q'$ ist also ein r. Q.-Körper über $\mathfrak{R}_0$. Aus der Maximalität von $\mathfrak{R}_q'$ folgt $\mathfrak{R}_q' \geqq \mathfrak{R}_q$ und aus der Maximalität von $\mathfrak{R}$ sowie $\mathfrak{R}_q$ ergibt sich $\mathfrak{R}_q' \leqq \mathfrak{R}$, $\mathfrak{R}_q' \leqq \mathfrak{R}_q$. Es ist also $\mathfrak{R}_q$ ein maximaler, nicht-normaler r. R.-Körper über $\mathfrak{R}_0$ und die Überlegungen von a) zeigen, daß bereits in $\mathfrak{R}_q$ ein Körper $\mathfrak{R}_0''$ der gewünschten Art enthalten ist.

d) Es sei jetzt $\mathfrak{R}_q = \mathfrak{N}_q$ über $\mathfrak{R}_0$ normal und $\mathfrak{R}_0 < \mathfrak{N}_q < \mathfrak{R}$. Nach Satz 2 gibt es in $\mathfrak{R}$ einen Körper $\mathfrak{R}_{q_1} > \mathfrak{N}_q$, bei dem $[\mathfrak{R}_{q_1} : \mathfrak{N}_q] = p$ eine ungerade Primzahl. Ist der zu $\mathfrak{R}_{q_1}$ über $\mathfrak{N}_q$ gehörige Normalkörper $\mathfrak{N}_{q_1}$ auch über $\mathfrak{R}_0$ normal, so ist $\mathfrak{R}_{q_1} = \mathfrak{R}_0''$ ein Körper des gesuchten Typs. Denn $\mathfrak{R}_{q_1}$ ist als r. R.-Körper vom Relativgrad p nicht einmal über $\mathfrak{N}_q$ normal, und es muß jeder zu $\mathfrak{R}_{q_1}$ über $\mathfrak{R}_0$ konjugierte Körper ein in $\mathfrak{N}_{q_1}$ enthaltener Oberkörper von $\mathfrak{N}_q$ vom Relativgrad p über $\mathfrak{N}_q$ sein. Die einzigen derartigen Körper sind aber die Konjugierten von $\mathfrak{R}_{q_1}$ über $\mathfrak{N}_q$, und daß unter diesen $\mathfrak{R}_{q_1}$ selbst der einzige reelle Körper ist, zeigt Hilfssatz 2. — Es bleibt also nur noch ein letzter schwierigster Fall übrig, nämlich: $\mathfrak{R}_0 < \mathfrak{R}_q = \mathfrak{N}_q < \mathfrak{R}$; zu dem Normalkörper $\mathfrak{N}_{q_1}$ der zu dem Körper $\mathfrak{R}_{q_1}$ über $\mathfrak{N}_q$ gehört, gibt es über $\mathfrak{R}_0$ mehrere konjugierte Körper $\mathfrak{N}_{q_1}, \mathfrak{N}_{q_2}, \ldots, \mathfrak{N}_{q_l}$. Hier erfordert die Konstruktion von $\mathfrak{R}_0''$ mehrere Schritte.

e) α) Nach § 2 enthält $\mathfrak{N}_{qi}$ einen über $\mathfrak{N}_q$ zyklischen Körper $\mathfrak{Z}_i$; wie in § 2 zerlegen wir $\mathfrak{Z}_i$ in zwei Faktoren: $\mathfrak{Z}_i = \mathfrak{Z}_{gi} \cdot \mathfrak{Z}_{ui}$ $[\mathfrak{Z}_{gi} : \mathfrak{N}_q] = 2^\varrho$, $[\mathfrak{Z}_{ui} : \mathfrak{N}_q] \not\equiv 0 (2)$ $(i = 1, \ldots, l)$. Natürlich sind die $\mathfrak{Z}_{gi}$ und ebenso die $\mathfrak{Z}_{ui}$ über $\mathfrak{R}_0$ konjugiert. Da $\mathfrak{R}_{q1}$ ein r. R.-Körper über $\mathfrak{N}_q$, ist $\mathfrak{Z}_{u1}$ nach Satz 3 der Vereinigungskörper

von $\mathfrak{R}_q$ und dem größten Unterkörper von $\mathfrak{K}_0(\varepsilon)$, der über $\mathfrak{K}_0$ ungeraden Grad besitzt. Das heißt aber: $\mathfrak{Z}_{u1}$ ist als Produkt zweier über $\mathfrak{K}_0$ normaler Körper selbst über $\mathfrak{K}_0$ normal, und wir haben daher: $\mathfrak{Z}_{u1}=\mathfrak{Z}_{u2}=\cdots=\mathfrak{Z}_{ul}=\mathfrak{Z}_u$. Weiter zeigen wir: Es ist auch $\mathfrak{Z}_{g1}=\mathfrak{Z}_{g2}=\cdots=\mathfrak{Z}_{gl}=\mathfrak{Z}_g$ und überdies $[\mathfrak{Z}_g:\mathfrak{R}_q]=2$. Es ist nämlich jedenfalls $\mathfrak{Z}_{g1}\cdot\mathfrak{Z}_{g2}\cdot\cdots\cdot\mathfrak{Z}_{gl}=\mathfrak{A}$ über $\mathfrak{R}_q$ Abelsch und $[\mathfrak{A}:\mathfrak{R}_q]=2^\varrho$ $(\varrho\geqq 1)$. Wäre nun $\varrho>1$, so wäre nach einer im wesentlichen schon in § 2 benützten Überlegung der größte reelle Unterkörper $\mathfrak{A}'$ von $\mathfrak{A}$ vom Grade $2^{\varrho-1}>1$ über $\mathfrak{R}_q$ und außerdem ein r. Q.-Körper über $\mathfrak{R}_q$. Das aber widerspricht der Maximaleigenschaft von $\mathfrak{R}_q$. Wir haben also: $[\mathfrak{A}:\mathfrak{R}_q]=2$ und das ist gerade die zu beweisende Behauptung.

β) Nach e, α) enthalten die über $\mathfrak{R}_q$ normalen Körper $\mathfrak{R}_{qi}$ alle ein und denselben größten über $\mathfrak{R}_q$ zyklischen Unterkörper $\mathfrak{Z}=\mathfrak{Z}_g\cdot\mathfrak{Z}_q$ und aus dem Beweis von Satz 3 in § 2 folgt: Es gibt einen r. Q.-Körper $\mathfrak{R}'_q$ über $\mathfrak{R}_q$, derart daß $\mathfrak{R}'_{qi}=\mathfrak{R}'_q\cdot\mathfrak{R}_{qi}=\mathfrak{R}'_q(\varepsilon,\alpha^{(i)})$ $\left(\varepsilon=\exp\left(\frac{2\pi i}{p}\right)\right)$, falls $\alpha^{(i)}$ jeweils in $\mathfrak{R}_{qi}$ so gewählt, daß $[\mathfrak{R}_q(\alpha^{(i)}):\mathfrak{R}_q]=p$. (Man beachte, daß die Existenz des kritischen r. Q.-Körpers bei Satz 3 allein auf Grund der speziellen Struktur von $\mathfrak{Z}$ bewiesen wurde.) Wir zeigen nun: $\mathfrak{R}'_q$ *ist normal über* $\mathfrak{K}_0$. — Es sei wie in § 2 $\mathfrak{E}_g$ der größte Unterkörper von $\mathfrak{K}_0(\varepsilon)$ mit $[\mathfrak{E}_g:\mathfrak{K}_0]=2^m$; $\mathfrak{E}'_g$ sei der größte reelle Unterkörper von $\mathfrak{E}_g$. Dann ist jedenfalls $\mathfrak{R}''_q=\mathfrak{E}'_g\cdot\mathfrak{R}_q$ als Produkt zweier über $\mathfrak{K}_0$ normaler Unterkörper selbst über $\mathfrak{K}_0$ normal und es ist $\mathfrak{R}_{q'}$ nach dem Beweis von Satz 3 der größte reelle Unterkörper von $\mathfrak{Z}_g\cdot\mathfrak{E}_g$, woraus wegen $[\mathfrak{E}_g:\mathfrak{E}'_g]=[\mathfrak{Z}_g:\mathfrak{R}_q]=2$ sofort zu schließen, daß entweder $[\mathfrak{Z}_g\cdot\mathfrak{E}_g:\mathfrak{R}''_q]=2$, $\mathfrak{R}''_q=\mathfrak{R}'_q$ oder $[\mathfrak{Z}_g\cdot\mathfrak{E}_g:\mathfrak{R}''_q]=4$, $[\mathfrak{R}'_q:\mathfrak{R}''_q]=2$. Im ersten Fall sind wir fertig. Im zweiten kann man so schließen. Es enthält $\mathfrak{Z}_g\cdot\mathfrak{E}_g$ drei über $\mathfrak{R}''_q$ quadratische Körper. Von ihnen sind zwei nämlich $\mathfrak{Z}_g\cdot\mathfrak{E}'_g$ und $\mathfrak{E}_g\cdot\mathfrak{R}''_q$, Produkte über $\mathfrak{K}_0$ normaler Körper und damit selbst normal; also muß auch $\mathfrak{R}'_q$, der dritte, über $\mathfrak{K}_0$ normal sein.

γ) Da $\mathfrak{K}_{q1}$ r. R.-Körper über $\mathfrak{K}_0$ und über $\mathfrak{R}_q$, existiert nach § 2 ein Binom x^p-b_1 $(b_1\in\mathfrak{R}'_q)$ mit der reellen Nullstelle β_1, derart daß $\mathfrak{K}_{q1}\subseteqq\mathfrak{R}'_q(\beta_1)$ und es gibt dann für *jede* Nullstelle β'_1 von x^p-b_1 in $\mathfrak{R}'_q(\beta'_1)$ einen Unterkörper von $\mathfrak{R}_{q1}$ vom Grade p über $\mathfrak{R}_q$. Es sei jetzt A_i ein Isomorphismus über $\mathfrak{K}_0$ von $\mathfrak{R}_{q1}$ auf $\mathfrak{R}_{qi}$, A'_i sei eine Fortsetzung von A_i, die $\mathfrak{R}'_{q1}=\mathfrak{R}'_q\cdot\mathfrak{R}_{q1}$ in $\mathfrak{R}'_{qi}=\mathfrak{R}'_q\cdot\mathfrak{R}_{qi}$ überführt. Bei A_i geht x^p-b_1 in ein Binom x^p-b_i $(b_i\in\mathfrak{R}'_q)$ über und aus Isomorphiegründen muß für jede Nullstelle β_i von x^p-b_i in $\mathfrak{R}'_q(\beta_i)$ ein Unterkörper $\mathfrak{K}_{qi}$ von $\mathfrak{R}_{qi}$ enthalten sein. Wählen wir für β_i speziell die reelle Nullstelle von x^p-b_i, so erweist sich $\mathfrak{K}_{qi}$ als r. R.-Körper über $\mathfrak{R}'_q$ und damit auch über $\mathfrak{K}_0$. Damit ist nach Hilfssatz 2 gleichzeitig gezeigt: Alle zu $\mathfrak{K}_{qi}$ über $\mathfrak{R}_q$ echt konjugierten Körper (und das sind alle von $\mathfrak{K}_{qi}$ verschiedenen Unterkörper von $\mathfrak{R}_{qi}$, die über $\mathfrak{R}_q$ den Grad p haben), sind komplex. Aus der letzten Bemerkung folgt sofort, die sämtlichen zu dem Körper $\mathfrak{K}''_0=\mathfrak{K}_{q1}\cdot\cdots\cdot\mathfrak{K}_{ql}$ über $\mathfrak{K}_0$ echt konjugierten Körper sind komplex. Denn bei jedem Automorphismus A des zu $\mathfrak{K}''_0$ über $\mathfrak{K}_0$ gehörigen Normalkörpers $\mathfrak{R}_{q1}\cdot\cdots\cdot\mathfrak{R}_{ql}$ vertauschen sich die $\mathfrak{R}_{qi}$ nur untereinander, und da $\mathfrak{R}_q$ über $\mathfrak{K}_0$ normal, muß bei A jeder der Körper $\mathfrak{K}_{qi}$ in einen Körper übergehen, der in

einem $\mathfrak{N}_{q^k}$ enthalten ist und über $\mathfrak{N}_q$ den Grad p besitzt. Im übrigen ist $\mathfrak{K}_0''$ nicht normal und ein r.R.-Körper über $\mathfrak{K}_0$, also in dem maximalen r.R.-Körper $\mathfrak{K}$ enthalten. Wir haben also auch im schwierigsten Fall die Konstruktion von $\mathfrak{K}_0'$ durchgeführt und damit den Beweis des Konjugiertensatzes abgeschlossen.

§4. Der Automorphismensatz. Ergänzungsbemerkungen

Der Konjugiertensatz für maximale r.R.-Körper wird ergänzt durch den Automorphismensatz:

Satz 8. Die Anzahl σ der Automorphismen eines maximalen r.R.-Körpers $\mathfrak{K}$ über seinem Grundkörper $\mathfrak{K}_0$ ist stets eine Potenz von 2: $\sigma = 2^m$ $(0 \leq m)$[14]*).*

Es sei $\mathfrak{N}$ der zu $\mathfrak{K}$ über $\mathfrak{K}_0$ gehörige Normalkörper und es werde für $\mathfrak{N} \geqq \mathfrak{L} \geqq \mathfrak{K}_0$ kurz $G_\mathfrak{L} = G(\mathfrak{N}:\mathfrak{L})$ gesetzt; $N_\mathfrak{L}(G_\mathfrak{K})$ sei der Normalisator von $G_\mathfrak{K}$ in $G_\mathfrak{L}$ $(\mathfrak{K} \geqq \mathfrak{L} \geqq \mathfrak{K}_0)$. Die zu bestimmende Automorphismenzahl ist dann gleich dem Index $[N_{\mathfrak{K}_0}(G_\mathfrak{K}):G_\mathfrak{K}]$ von $G_\mathfrak{K}$ in $N_{\mathfrak{K}_0}(G_\mathfrak{K})$. — Für $\mathfrak{K} = \mathfrak{K}_0$ wird $G_\mathfrak{K} = N_{\mathfrak{K}_0}(G_\mathfrak{K}) = G_{\mathfrak{K}_0}$, $\sigma = 1 = 2^0$. Wir dürfen also die Induktionsvoraussetzung machen, daß für $\mathfrak{K} \geqq \mathfrak{K}_1 > \mathfrak{K}_0$ bereits $[N_{\mathfrak{K}_1}(G_\mathfrak{K}):G_\mathfrak{K}] = 2^{m_1}$ bewiesen sei. Es sei dabei $\mathfrak{K}_1$ nach Satz 2 insbesondere so gewählt, daß $[\mathfrak{K}_1:\mathfrak{K}_0]$ Primzahl. Dann unterscheiden wir zwei Fälle:

a) $[\mathfrak{K}_1:\mathfrak{K}_0] = p \neq 2$. Hier besitzt der Körper $\mathfrak{K}_1$, der ja nach Hilfssatz 2 nicht über $\mathfrak{K}_0$ normal ist, nur den identischen Automorphismus, und es wird daher $N_{\mathfrak{K}_0}(G_\mathfrak{K}) = N_{\mathfrak{K}_1}(G_\mathfrak{K})$, $\sigma = [N_{\mathfrak{K}_1}(G_\mathfrak{K}):G_\mathfrak{K}] = 2^{m_1}$.

b) $[\mathfrak{K}_1:\mathfrak{K}_0] = 2$. Hier gibt es einen einzigen echten Automorphismus A_1 von $\mathfrak{K}_1$ über $\mathfrak{K}_0$. Wir haben daher $[N_{\mathfrak{K}_0}(G_\mathfrak{K}):N_{\mathfrak{K}_1}(G_\mathfrak{K})] = 2$ bzw. $= 1$, je nachdem ob sich A_1 zu einem Automorphismus A von $\mathfrak{K}$ über $\mathfrak{K}_0$ fortsetzen läßt oder nicht. Mit der Voraussetzung $[N_{\mathfrak{K}_1}(G_\mathfrak{K}):G_\mathfrak{K}] = 2^{m_1}$ erhalten wir somit für σ die beiden Möglichkeiten $\sigma = 2^{m_1+1}$ bzw. $\sigma = 2^{m_1}$. — Nennen wir $P(x)$ ein erzeugendes Polynom von $\mathfrak{K}$ über $\mathfrak{K}_0$, falls $\mathfrak{K}$ zum Restklassenkörper $\mathfrak{K}_0[x]/P(x)$ isomorph ist, so ist die Anzahl der Automorphismen von $\mathfrak{K}$ über $\mathfrak{K}_0$ gleich der Anzahl aller der Nullstellen α eines erzeugenden Polynoms $Q(x)$, für die $\mathfrak{K} = \mathfrak{K}_0(\alpha)$ wird. Aus dieser Bemerkung ergibt sich mühelos, daß Satz 7 und Satz 8 zusammengenommen gleichwertig sind mit dem folgenden Theorem:

Satz 9. Ist $P(x)$ ein erzeugendes Polynom des maximalen r.R.-Körpers $\mathfrak{K}$, so ist die Anzahl der reellen Nullstellen von $P(x)$ eine Potenz von 2 und es wird $\mathfrak{K} = \mathfrak{K}_0(\alpha)$ für jede reelle Nullstelle α von $P(x)$.

---

[14]) Über den Exponenten m läßt sich vermutlich keine genauere Aussage machen, abgesehen davon, daß trivialerweise $m = 0$, wenn $[\mathfrak{K}:\mathfrak{K}_0] \not\equiv 0\,(2)$ und $m \geq 1$, wenn $[\mathfrak{K}:\mathfrak{K}_0] \equiv 0\,(2)$. Sicher ist *nicht* allgemein 2^m gleich der höchsten in $[\mathfrak{K}:\mathfrak{K}_0]$ aufgehenden Potenz von 2. Beispiel: $\mathfrak{K}_0$ der Körper der rationalen Zahlen, $\alpha = +\sqrt{1+\sqrt{3}}$, $\beta = +\sqrt{1-\sqrt{3}}$. Hier ist, da β komplex, $\mathfrak{K}_0(\alpha) = \mathfrak{K}$ der maximale r.R.-Körper des Normalkörpers $\mathfrak{N} = \mathfrak{K}_0(\alpha, \beta)$, — ($[\mathfrak{K}_0(\alpha, \beta):\mathfrak{K}_0] = 8$ beachten!) —, und es ist $[\mathfrak{K}_0(\alpha):\mathfrak{K}_0] = 4$. Aber $\mathfrak{K}$ hat über $\mathfrak{K}_0$ nur zwei Automorphismen (charakterisiert durch $\alpha \to \alpha$ und $\alpha \to -\alpha$) und daneben zwei Isomorphismen auf einen echt konjugierten Körper (charakterisiert durch $\alpha \to \beta$ und $\alpha \to -\beta$).

Mit Satz 9 ist der Anschluß an die ältere Ausdrucksweise hergestellt, die lieber mit den erzeugenden Polynomen bzw. definierenden Gleichungen als mit den Körpern operiert. Man beachte aber, daß die Beweise durch die Hereinziehung der erzeugenden Polynome nur unnötig schwerfällig würden. — Aus Satz 9 folgt insbesondere: Ein erzeugendes Polynom $P(x)$ des maximalen r. R.-Körpers $\mathfrak{R}$ besitzt dann und nur dann genau *eine* reelle Nullstelle, wenn $[\mathfrak{R}:\mathfrak{R}_0]$ ungerade ist. — Bei ungeradem Grad kann man übrigens die Forderung der Maximalität etwas abschwächen. Wir wollen kurz sagen, der Körper $\mathfrak{R}$ besitze (über seinem Grundkörper $\mathfrak{R}_0$) eine *Primkette*, wenn eine Körperfolge $\mathfrak{R}_0 \subset \mathfrak{R}_1 \subset \cdots \subset \mathfrak{R}_m = \mathfrak{R}$ existiert, bei der alle Relativgrade $[\mathfrak{R}_i:\mathfrak{R}_{i-1}]$ Primzahlen sind. Nach Satz 2 hat jeder maximale r. R.-Körper Primketten, aber natürlich gibt es auch r. R.-Körper mit Primkette, die nicht maximal sind. Eine nicht triviale Ergänzung von Satz 9 ist daher Satz 10:

Satz 10. Besitzt der r. R.-Körper $\mathfrak{R}$ eine Primkette und ist außerdem $[\mathfrak{R}:\mathfrak{R}_0]$ ungerade, so hat jedes erzeugende Polynom $P(x)$ von $\mathfrak{R}$ genau eine reelle Nullstelle.

Zum Beweise von Satz 10 beachte man: a) Die Behauptung läßt sich auch so formulieren: Alle echt konjugierten Körper von $\mathfrak{R}$ sind komplex und es gibt keinen echten Automorphismusvon $\mathfrak{R}$ über $\mathfrak{R}_0$. b) Unsere Behauptung ist nach Hilfssatz 2 richtig, wenn $[\mathfrak{R}:\mathfrak{R}_0]$ Primzahl. Ist daher $\mathfrak{R}_0 \subset \mathfrak{R}_1 \subset \cdots \subset \mathfrak{R}_m = \mathfrak{R}$ eine Primkette von $\mathfrak{R}$ über $\mathfrak{R}_0$, so dürfen wir nach Induktionsprinzipien annehmen, daß sie [in der Fassung a)] sowohl für $\mathfrak{R}_1$ über $\mathfrak{R}_0$ als auch für $\mathfrak{R}$ über $\mathfrak{R}_1$ zutrifft. Daraus ergibt sich aber sofort die Gültigkeit der Behauptung für $\mathfrak{R}$ über $\mathfrak{R}_0$. Satz 10 ist eine unmittelbare Verallgemeinerung von Hilfssatz 2, der den Fall erledigt, daß eine nur aus zwei Gliedern ($\mathfrak{R}_0$ und $\mathfrak{R}$) bestehende Primkette existiert. Dagegen läßt sich Hilfssatz 2 nur mit einigem Zwang als Spezialfall von Satz 9 auffassen — ein r. R.-Körper vom Primzahlgrad über dem Grundkörper braucht ja nicht maximal zu sein. Da aber der direkte Beweis von Hilfssatz 2 äußerst einfach ist[15]) brauchen wir auf diesen Punkt nicht näher einzugehen. — Natürlich liegt es nahe, zu fragen: Gilt etwa Satz 10 auch dann noch, wenn man auf die Existenz einer Primkette verzichtet und nur $[\mathfrak{R}:\mathfrak{R}_0] \not\equiv 0(2)$ fordert? Ehe man hier eine unmittelbare Entscheidung anstrebt, wird man am besten zuerst ein Beispiel für einen r. R.-Körper $\mathfrak{R}$ suchen, der keine Primkette besitzt, obwohl der Grad $[\mathfrak{R}:\mathfrak{R}_0]$ ungerade ist. Ein solches Beispiel dürfte wesentlich mühsamer zu bilden sein, als das in § 1 angegebene Beispiel für einen r. R.-Körper ohne Primkette

[15]) Es verdient Erwähnung, daß beim Beweise von Hilfssatz 2 keine speziellen Eigenschaften der Binome und vor allem nicht die Einheitswurzeln gebraucht werden. Man überlegt so: Es sei $P(x)$ ein irreduzibles Polynom über dem reellen Körper $\mathfrak{R}_0$ vom Primzahlgrad $p \neq 2$ mit auflösbarer Gruppe und es sei $P(x)$ noch über dem reellen Körper $\mathfrak{R}_1 \geqq \mathfrak{R}_0$ irreduzibel, während $P(x)$ über $\mathfrak{R}_1(\beta_0)$ reduzibel wird, wobei β_0 über $\mathfrak{R}_1$ reelle Nullstelle des irreduzibeln Polynoms $Q(x)$ vom Primzahlgrade q ist und $Q(x)$ außer β_0 nur komplexe Nullstellen $\beta_1, \ldots, \beta_{q-1}$ besitzt. Dann spaltet $P(x)$ wegen der Auflösbarkeit seiner Gruppe über $\mathfrak{R}_1(\beta_0)$ einen reellen Linearfaktor ab und man hat daher: $p = q$; $\alpha_0 = \sum_{k=0,\ldots,p-1} c_k \cdot \beta_0^k = c(\beta_0)$ $(c_k \in \mathfrak{R}_1)$, während $\alpha_i = c(\beta_i)$ $(i = 1, \ldots, q-1)$ die übrigen Nullstellen von $P(x)$ sind. Wäre nun etwa α_1 reell, $\alpha_1 = \bar{\alpha}_1$ und $\beta_1 = \bar{\beta}_i$ $(i \neq 1!)$, so hätte man $\alpha_1 = c(\beta_1) = \bar{\alpha}_1 = c(\bar{\beta}_1) = c(\beta_i) = \alpha_i$, was unmöglich.

mit $[\mathfrak{K}:\mathfrak{K}_0]=4$. Auch besteht immerhin die entfernte Möglichkeit, daß man hier auf eine Ausnahmestellung der Primzahl 2 stößt, wie sie bei unseren bisherigen Betrachtungen schon öfters hervortrat. Hier können wir diese Fragen auf sich beruhen lassen. Denn ein Vergleich des Beweises von Satz 10 mit dem Beweis von Satz 7 zeigt, daß man durch die Forderung $[\mathfrak{K}:\mathfrak{K}_0] \not\equiv 0(2)$ den wirklich interessanten, in der allgemeinen Theorie der r.R.-Körper auftretenden Schwierigkeiten aus dem Wege geht.

Dagegen soll kurz durch ein Beispiel gezeigt werden, daß ohne einschränkende Annahmen über $[\mathfrak{K}:\mathfrak{K}_0]$ Satz 7 und 9 *nur* für *maximale* r.R.-Körper ausgesprochen werden können. Wir betrachten über dem Körper $\mathfrak{K}_0$ der rationalen Zahlen das Polynom $p(x)=x^6+6\cdot x^3-63\cdot x^2+9$. Über $\mathfrak{K}_1=\mathfrak{K}_0(\sqrt{7})$ zerfällt $p(x)$ in die Faktoren $p_1(x)=x^3+3\cdot\sqrt{7}\cdot x+3$, $p_2(x)=x^3-3\cdot\sqrt{7}\cdot x+3$. Da in der Hauptordnung aller ganzen algebraischen Zahlen des Körpers $\mathfrak{K}_1$ das Hauptideal (3) keinen Primidealfaktor in höherer als erster Potenz enthält, sind $p_1(x)$ und $p_2(x)$ nach SCHÖNEMANN-EISENSTEIN über $\mathfrak{K}_1$ irreduzibel, und daraus folgt sofort die Irreduzibilität von $p(x)$ über $\mathfrak{K}_0$. Das Polynom dritten Grades $p_1(x)$ besitzt genau *eine* reelle Nullstelle ϱ, und es ist daher $\mathfrak{K}_0(\varrho,\sqrt{7})$ r.R.-Körper über $\mathfrak{K}_1=\mathfrak{K}_0(\sqrt{7})$ und damit auch über $\mathfrak{K}_0$. Ferner hat man: $[\mathfrak{K}_0(\varrho):\mathfrak{K}_0]=6$, $[\mathfrak{K}_1(\varrho):\mathfrak{K}_0]=[\mathfrak{K}_1(\varrho):\mathfrak{K}_1]\cdot[\mathfrak{K}_1:\mathfrak{K}_0]=3\cdot 2=6$, also $\mathfrak{K}_1(\varrho)=\mathfrak{K}_0(\varrho,\sqrt{7})=\mathfrak{K}_0(\varrho)$. Es ist also $\mathfrak{K}_0(\varrho)$ ein r.R.-Körper über $\mathfrak{K}_0$ mit Primkette. Aber $p_2(x)$ hat drei reelle Nullstellen $\sigma_1,\sigma_2,\sigma_3$ und es sind daher nach Hilfssatz 2 die zu $\mathfrak{K}_0(\varrho)$ über $\mathfrak{K}_0$ konjugierten reellen Körper $\mathfrak{K}_0(\sigma_i)$ keine r.R.-Körper über $\mathfrak{K}_1$ und damit auch keine über $\mathfrak{K}_0$. Wir können daher feststellen:

Es gibt r.R.-Körper mit Primkette, die echt konjugierte reelle Körper besitzen, und zwar sogar solche, die ihrerseits keine r.R.-Körper über dem gegebenen Grundkörper sind[16]).

Der letzte Teil dieses Resultats ist besonders wichtig. Denn der Beweis des Konjugiertensatzes 7 wäre trivial gewesen, wenn man hätte zeigen können: Jeder zu einem r.R.-Körper konjugierte Körper muß selbst ein r.R.-Körper sein. — Nachdem aber diese Vermutung widerlegt ist, hat man den Eindruck, daß der Beweis von Satz 7 kaum wesentlich abgekürzt werden kann, und daß vor allem ein sorgfältiger Induktionsschluß nicht zu vermeiden ist.

[16]) Daß in unserem Beispiel die Anzahl der reellen Nullstellen von $p(x)$ $4=2^2$ ist uninteressant, da es ja auf die Zahl *der* reellen Nullstellen ankommt, die den Körper $\mathfrak{K}$ erzeugen. Es wäre übrigens wohl nicht schwer, ein Beispiel zu konstruieren, in dem die Zahl aller reellen Nullstellen keine zweier Potenz würde. Man müßte dazu über dem Körper $\mathfrak{K}_0$ der rationalen Zahlen das irreduzible Polynom $p(x)$ vom Grade 10 so bestimmen, daß $p(x)$ über $\mathfrak{K}_0(\sqrt{d})$ $(d>0)$ in zwei irreduzible Faktoren $p_1(x)$, $p_2(x)$ des Grades 5 zerfällt, wobei $p_1(x)$ eine, $p_2(x)$ dagegen fünf reelle Nullstellen besitzt. [Eine rechnerische Erschwerung für die Konstruktion liegt nur darin, daß man hier $p_1(x)$ und $p_2(x)$ nicht trinomisch annehmen kann.]

Bonn, Mathematisches Institut der Universität

(Eingegangen am 24. Oktober 1955)

NAUMANN, H.
Math. Zeitschr. Bd. 65, S. 91—103 (1956)

Beliebige konvexe Polytope als Schnitte und Projektionen höherdimensionaler Würfel, Simplizes und Maßpolytope*)

Von
HERBERT NAUMANN

In dieser Arbeit wird gezeigt, daß sich jedes konvexe Polytop als Schnitt eines höherdimensionalen Würfels und als Projektion eines höherdimensionalen Maßpolytops auffassen läßt. Für Mittelpunktspolytope ist speziell eine Darstellung als Mittelpunktsschnitt bzw. senkrechte Parallelprojektion möglich. Ebenfalls wird die Darstellbarkeit beliebiger konvexer Polytope als Schnitt und Projektion höherdimensionaler regulärer Simplizes gezeigt. Es werden Abschätzungen für die Dimension des höherdimensionalen Würfels usw. gegeben.

1. Sei zunächst in einem n-dimensionalen euklidischen Raum ein *konvexes Polytop mit Mittelpunkt* gegeben. Wir wollen voraussetzen, daß ein Polytop mindestens einen inneren Punkt besitzen soll. Folglich ist dann der diametral gelegene Punkt ebenfalls ein innerer Punkt, und damit wegen der Konvexität auch der Mittelpunkt des Polytops innerer Punkt. Der Mittelpunkt liege im Koordinatenursprung. Dann lassen sich die $2m$ extremen Stützen des gegebenen Polytops analytisch in der Normalform

$$(1) \qquad \left.\begin{aligned} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n &= +1 \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n &= -1 \end{aligned}\right\} \quad (i = 1, 2, \ldots, m)$$

darstellen. Die Punkte des von den extremen Stützen eingeschlossenen Raums genügen ersichtlich der Beziehung

$$(2) \qquad \max_{i=1}^{m} \left| \sum_{k=1}^{n} a_{ik} x_k \right| \leq 1,$$

wobei das Gleichheitszeichen genau auf dem Rande gilt. Es läßt sich also jedes konvexe Mittelpunktspolytop (genauer: sein Rand) durch eine Gleichung

$$(3) \qquad \max_{i=1}^{m} \left| \sum_{k=1}^{n} a_{ik} x_k \right| = 1$$

darstellen. Umgekehrt stellt eine Gleichung der Gestalt (3) genau dann den Rand eines konvexen Mittelpunktspolytops dar, wenn das Gebilde nur Punkte

*) Für Unterstützung dieser Arbeit sage ich der Deutschen Forschungsgemeinschaft meinen aufrichtigen Dank.

im Endlichen besitzt, d.h. wenn nach Einführung homogener Koordinaten $\bar{x}_k = \bar{x}_0 x_k$ die Gleichung

$$\max_{i=1}^{m} \left| \sum_{k=1}^{n} a_{ik} \bar{x}_k \right| = 0 \tag{4}$$

keine nichttriviale Lösung besitzt. Da dies äquivalent damit ist, daß das homogene Gleichungssystem

$$\sum_{k=1}^{n} a_{ik} \bar{x}_k = 0 \qquad (i = 1, \ldots, m) \tag{5}$$

keine nichttriviale Lösung besitzt, so ergibt sich als notwendige und hinreichende Bedingung dafür, daß das Gebilde (3) beschränkt ist, daß die Matrix $((a_{ik}))$ vom Range $\varrho = n$ ist. Konvexes Polytop ist sodann das von (3) berandete Gebilde als beschränkter Durchschnitt endlich vieler Stützstreifen.

Sei nun ein *beliebiges konvexes Polytop* des R_n gegeben. Die m extremen Stützen seien

$$\sum_{k=1}^{n} a_{ik} x_k = a_i, \qquad i = 1, \ldots, m. \tag{6}$$

Dann läßt sich zu jedem a_i ein $\bar{a}_i$ finden mit der Eigenschaft, daß das Polytop ganz zwischen der Hyperebene $\sum_{k=1}^{n} a_{ik} x_k = a_i$ und der Hyperebene $\sum_{k=1}^{n} a_{ik} x_k = \bar{a}_i$ liegt. Nach Subtraktion von $\frac{a_i + \bar{a}_i}{2}$ von beiden Seiten beider Gleichungen und anschließende Division durch $\frac{a_i - \bar{a}_i}{2}$ erhalten wir, daß sich jede extreme Stütze $\sum_{k=1}^{n} a_{ik} x_k = a_i$ derart in der Form $a'_i + \sum_{k=1}^{n} a'_{ik} x_k = 1$ schreiben läßt, daß das Polytop ganz zwischen $a'_i + \sum_{k=1}^{n} a'_{ik} x_k = +1$ und $a'_i + \sum_{k=1}^{n} a'_{ik} x_k = -1$ liegt. Ganz entsprechend wie bei den Mittelpunktspolytopen ergibt sich daraus

Satz 1.1. *Jedes konvexe Polytop des R_n mit m extremen Stützen läßt sich in der Form*

$$\max_{i=1}^{m} \left| a_i + \sum_{k=1}^{n} a_{ik} x_k \right| = 1 \tag{7}$$

darstellen.

Es bilden auch umgekehrt die Punkte, die einer Gleichung der Gestalt (7) genügen, genau dann den Rand eines konvexen Polytops, wenn es einen Punkt $(\xi_1, \xi_2, \ldots, \xi_n)$ gibt, für den $\max_{i=1}^{m} \left| a_i + \sum_{k=1}^{n} a_{ik} \xi_k \right| < 1$ ist und die Matrix $((a_{ik}))$ vom Range $\varrho = n$ ist. Die erstere Bedingung brauchte im Falle der Mittelpunktspolytope nicht besonders angegeben werden, da in der Darstellung (3) der Punkt $(0, 0, \ldots, 0)$ stets ein innerer Punkt ist. Die zweite Bedingung ergibt sich entsprechend wie bei den Mittelpunktspolytopen.

2. Eine ganz analoge Darstellung ergibt sich bei Verwendung von Hyperebenenkoordinaten.

Seien die p Ecken eines konvexen Polytops durch ihre Punktkoordinaten

$$(a_{i1}, a_{i2}, \ldots, a_{in}) \qquad i = 1, \ldots, p \tag{1}$$

gegeben und der Punkt $(0, 0, \ldots, 0)$ ein innerer Punkt des Polytops. Durch die spezielle Polarkorrelation

$$x_k^* = u_k, \quad u_k^* = x_k \qquad (k = 1, \ldots, n), \tag{2}$$

wobei die u_k als Koordinaten der Hyperebene

$$\sum_{k=1}^{n} u_k x_k = 1 \tag{3}$$

aufzufassen sind, geht das Polytop wieder in ein konvexes Polytop über, das den Punkt $(0, 0, \ldots, 0)$ als inneren Punkt besitzt. Die Ecken des ursprünglichen Polytops sind hierbei in die extremen Stützen

$$\sum_{k=1}^{n} a_{ik} x_k^* = 1, \qquad i = 1, \ldots, p \tag{4}$$

des transformierten Polytops übergegangen. Nach Satz 1.1 läßt sich dieses in der Form

$$\max_{i=1}^{p} \left| b_i + \sum_{k=1}^{n} b_{ik} x_k^* \right| = 1 \tag{5}$$

darstellen. Macht man nun die Korrelation (2) wieder rückgängig, so ergibt sich

$$\max_{i=1}^{p} \left| b_i + \sum_{k=1}^{n} b_{ik} u_k \right| = 1 \tag{6}$$

als Darstellung des ursprünglich gegebenen Polytops in Hyperebenenkoordinaten, und zwar genügen der Gl. (6) genau die Hyperebenen, die das Polytop umhüllen und mindestens einen Punkt mit seinem Rand gemeinsam haben.

Satz 2.1. *Jedes konvexe Polytop des R_n mit p Ecken läßt sich in der Form*

$$\max_{i=1}^{p} \left| b_i + \sum_{k=1}^{n} b_{ik} u_k \right| = 1 \tag{7}$$

darstellen.

Es ergibt sich unmittelbar, daß sich jedes konvexe Mittelpunktspolytop mit $2p$ Ecken in der Form

$$\max_{i=1}^{p} \left| \sum_{k=1}^{n} b_{ik} u_k \right| = 1 \tag{8}$$

darstellen läßt, wo

$$\left.\begin{array}{c} (b_{i1}, b_{i2}, \ldots, b_{in}) \\ (-b_{i1}, -b_{i2}, \ldots, -b_{in}) \end{array}\right\} \qquad i = 1, \ldots, p$$

die Ecken des Polytops in Punktkoordinaten sind.

3. Wir betrachten nun eine spezielle Klasse konvexer Polytope, die hier „Hauptpolytope" genannt seien.

Definition. Ein Hauptpolytop ist ein Polytop, welches bei geeigneter Lage im Koordinatensystem folgende Symmetrieeigenschaften besitzt:

Mit $(a_1, a_2, \ldots, a_{n-1}, a_n)$ sind auch $(a_2, a_3, \ldots, a_n, a_1)$ und $(-a_1, a_2, \ldots, a_{n-1}, a_n)$ Punkte des Polytops.

Offenbar ist diese Bedingung zu jeder der beiden folgenden Bedingungen äquivalent:

a) Mit $b_1 x_1 + b_2 x_2 + \cdots + b_{n-1} x_{n-1} + b_n x_n = 1$ sind auch $b_2 x_1 + b_3 x_2 + \cdots + b_n x_{n-1} + b_1 x_n = 1$ und $-b_1 x_1 + b_2 x_2 + \cdots + b_{n-1} x_{n-1} + b_n x_n = 1$ extreme Stützen des Polytops.

b) Mit $(a_1, a_2, \ldots, a_{n-1}, a_n)$ sind auch $(a_2, a_3, \ldots, a_n, a_1)$ und $(-a_1, a_2, \ldots, a_{n-1}, a_n)$ Eckpunkte des Polytops.

Da die beiden letzten Bedingungen dual zueinander sind, so ist das zu einem Hauptpolytop polare Polytop wieder ein Hauptpolytop.

Beispiele für Hauptpolytope sind alle regulären Polytope mit Ausnahme der Simplizes und der regulären $(4k+1)$-, $(4k+2)$- und $(4k+3)$-Ecke, weitere Beispiele sind Kubooktaeder, Rhombendodekaeder, Triakisoktaeder, Tetrakishexaeder usw.

Bringt man ein Hauptpolytop mit $2m$ extremen Stützen und $2p$ Ecken nach dem in § 1 bzw. § 2 beschriebenen Verfahren auf die Gestalt (1.3) bzw. (2.8), so genügen dann die Koeffizienten a_{ik} bzw. b_{ik} den „Orthogonalitätsrelationen"

$$\sum_{i=1}^{m} a_{ik} a_{il} = \begin{cases} c & \text{für } k = l \\ 0 & \text{für } k \neq l \end{cases} \qquad (k, l = 1, \ldots, n) \tag{1}$$

bzw.

$$\sum_{i=1}^{p} b_{ik} b_{il} = \begin{cases} c' & \text{für } k = l \\ 0 & \text{für } k \neq l \end{cases} \qquad (k, l = 1, \ldots, n). \tag{2}$$

Wir begnügen uns mit dem Nachweis von (1), da der Nachweis von (2) analog verläuft. Die Koeffizientenmatrix des gegebenen Hauptpolytops ist

$$\mathfrak{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Nach der Bedingung a) für Hauptpolytope müssen die Polytope mit den Koeffizientenmatrizen

$$\mathfrak{A}_1 = \begin{pmatrix} a_{12} & a_{13} & \cdots & a_{1n} & a_{11} \\ a_{22} & a_{23} & \cdots & a_{2n} & a_{21} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m2} & a_{m3} & \cdots & a_{mn} & a_{m1} \end{pmatrix}, \qquad \mathfrak{A}_2 = \begin{pmatrix} -a_{11} & a_{12} & \cdots & a_{1n} \\ -a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

mit dem gegebenen Polytop identisch sein. Da die Hyperebenen $\left|\sum_{k=1}^{n} a_{ik} x_k\right| = 1$ $(i = 1, \ldots, m)$ umkehrbar eindeutig den $2m$ extremen Stützen des Polytops zugeordnet sind, so entstehen also die Zeilen von $\mathfrak{A}_1$ und die Zeilen von $\mathfrak{A}_2$ durch eine Umordnung der Zeilen von $\mathfrak{A}$ unter eventueller Multiplikation aller Glieder einer Zeile mit -1. Folglich ist die Quadratsumme der Glieder einer Spalte für alle Spalten konstant und $\sum_{i=1}^{m} a_{i1} a_{il} = \sum_{i=1}^{m} (-a_{i1}) a_{il} = 0$ für $l = 2, \ldots, n$, woraus durch wiederholte Anwendung der Eigenschaften eines Hauptpolytops schließlich folgt, daß die Spaltenvektoren aus $\mathfrak{A}$ paarweise orthogonal zueinander sind, womit der Beweis unserer Behauptung vollständig erbracht ist.

4. In diesem Paragraphen zeigen wir:

Satz 4.1. *Jedes Mittelpunktspolytop des R_n, das eine Darstellung*

$$\max_{i=1}^{m} \left|\sum_{k=1}^{n} a_{ik} x_k\right| = 1 \quad \text{mit} \quad \sum_{i=1}^{m} a_{ik} a_{il} = \begin{cases} c & \text{für } k = l \\ 0 & \text{für } k \neq l \end{cases} \quad (k, l = 1, \ldots, n) \tag{1}$$

besitzt, läßt sich als Mittelpunktsschnitt eines m-dimensionalen Würfels mit einem n-dimensionalen linearen Raum darstellen.

Satz 4.2. *Jedes Mittelpunktspolytop des R_n, das eine Darstellung*

$$\max_{i=1}^{p} \left|\sum_{k=1}^{n} b_{ik} u_k\right| = 1 \quad \text{mit} \quad \sum_{i=1}^{p} b_{ik} b_{il} = \begin{cases} c' & \text{für } k = l \\ 0 & \text{für } k \neq l \end{cases} \quad (k, l = 1, \ldots, n) \tag{2}$$

besitzt, läßt sich als senkrechte Parallelprojektion eines p-dimensionalen Maßpolytops auf einen n-dimensionalen linearen Raum darstellen.

Mit den Ergebnissen von § 3 ergibt sich hieraus als unmittelbare Folgerung:

Satz 4.3. *Jedes konvexe Hauptpolytop, das von $2m$ extremen Stützen begrenzt wird und $2p$ Ecken besitzt, ist als Mittelpunktsschnitt des m-dimensionalen Würfels und als senkrechte Parallelprojektion des p-dimensionalen Maßpolytops darstellbar.*

Zum Beweis von Satz 4.1 setzen wir

$$d_{ik} = \frac{a_{ik}}{\sqrt{c}} \quad (i = 1, \ldots, m, \; k = 1, \ldots, n). \tag{3}$$

Die Spaltenvektoren der Matrix $((d_{ik}))$ bilden dann n normierte Einheitsvektoren mit m Komponenten, es läßt sich also die Matrix $((d_{ik}))$ durch Hinzufügen von $m - n$ weiteren Spalten zu einer orthogonalen m-reihigen Matrix erweitern. Der m-dimensionale Würfel der Kantenlänge $\frac{2}{\sqrt{c}}$ läßt sich bekanntlich in der Form

$$\max_{i=1}^{m} \left|\sqrt{c}\, x_i'\right| = 1 \tag{4}$$

und nach Anwendung der orthogonalen Substitution

$$x_i' = \sum_{k=1}^{m} d_{ik} x_k \qquad (i = 1, \ldots, m) \tag{5}$$

folglich auch in der Form

$$\max_{i=1}^{m} \sum_{k=1}^{m} \left| c\, d_{ik} x_k \right| = 1 \tag{6}$$

darstellen. Durch Schnitt dieses Würfels mit dem n-dimensionalen linearen Raum $x_{n+1} = x_{n+2} = \cdots = x_m = 0$ ergibt sich mittels (3) in der Tat das gegebene Polytop.

Der Beweis von Satz 4.2 folgt hieraus durch Dualisierung.

5. Aus dem Satz 4.3 für konvexe Hauptpolytope leiten wir nun einen Beweis des Satzes her, daß sich jedes beliebige konvexe Polytop als Schnitt eines höherdimensionalen Würfels und als Projektion eines höherdimensionalen Maßpolytops darstellen läßt, und geben gleichzeitig eine Abschätzung für die Dimensionszahl des höherdimensionalen Würfels bzw. Maßpolytops.

Wir betrachten im R_n irgendeine Hyperebene

$$\sum_{i=1}^{n} a_i x_i = 1, \tag{1}$$

wobei die Koeffizienten a_i ihrem absoluten Betrag nach paarweise untereinander und von Null verschieden seien. Durch wiederholte Anwendung der Symmetrieoperationen, die Hauptpolytope invariant lassen, wird von der Hyperebene (1) ein Hauptpolytop mit den extremen Stützen

$$\sum_{i=1}^{n} (-1)^{\mu_i} a_{i+\nu_i} x_i = 1 \qquad (\mu_i = 0, 1,\ \nu_i = 0, \ldots, n-1;\ \text{Indizes mod } n) \tag{2}$$

erzeugt; die Anzahl dieser extremen Stützen beträgt nach der Voraussetzung über die a_i genau $2^n \cdot n$.

Bringt man nun im Inneren eines der $(n-1)$-dimensionalen Begrenzungspolytope des soeben konstruierten Hauptpolytops ein beliebiges konvexes Polytop des R_{n-1}, welches p Ecken besitzen möge, unter (eventuell nach entsprechender Streckung des Hauptpolytops), überträgt dieses durch Anwendung der erwähnten Symmetrieoperationen auf die $2^n \cdot n - 1$ übrigen Begrenzungspolytope des Hauptpolytops und bildet die konvexe Hülle des $(n-1)$-dimensionalen Polytops und seiner $2^n \cdot n - 1$ Bilder, so entsteht ein (dem ursprünglichen Hauptpolytop einbeschriebenes) konvexes Hauptpolytop mit $2^n \cdot n \cdot p$ Ecken, das unter seinen $(n-1)$-dimensionalen Begrenzungspolytopen $2^n \cdot n$ dem vorgegebenen konvexen Polytop des R_{n-1} kongruente Polytope enthält.

Nach Satz 4.3 ist das zuletzt gefundene Hauptpolytop als Projektion des $(2^{n-1} \cdot n \cdot p)$-dimensionalen Maßpolytops darstellbar. Legt man nun durch eines der oben erwähnten Begrenzungspolytope mit p Ecken einen $(n-1)$-

dimensionalen Raum als Projektionsraum, so wird das Hauptpolytop von einem außerhalb und genügend nahe über dem zugehörigen Begrenzungspolytop gelegenen Projektionszentrum auf dieses Begrenzungspolytop projiziert. Da die Projektion einer Projektion wieder eine Projektion liefert, so ist damit, wenn man n durch $n+1$ ersetzt, erhalten:

Satz 5.1. *Jedes beliebige konvexe Polytop des R_n mit p Ecken ist als Projektion eines $(2^n \cdot (n+1) \cdot p)$-dimensionalen Maßpolytops darstellbar.*

Das Dualitätsprinzip liefert hierzu:

Satz 5.2. *Jedes beliebige konvexe Polytop des R_n mit m extremen Stützen ist als Schnitt eines $(2^n \cdot (n+1) \cdot m)$-dimensionalen Würfels darstellbar.*

6. Wir entwickeln nun auf einem anderen Wege eine Abschätzung für den Würfelschnittsatz und den hierzu dualen Satz speziell für Mittelpunktspolytope.

Sei ein beliebiges konvexes Mittelpunktspolytop

$$\max_{i=1}^{p} \left| \sum_{k=1}^{n} b_{ik} u_k \right| = 1 \tag{1}$$

gegeben. Wir setzen jetzt

$$\mathfrak{r}_i = \begin{pmatrix} b_{i1} \\ b_{i2} \\ \vdots \\ b_{in} \end{pmatrix} \quad (i = 1, \ldots, p), \tag{2}$$

$$\mathfrak{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{3}$$

und führen den Richtungsvektor

$$\mathfrak{c} = \begin{pmatrix} \cos \alpha_1 \\ \cos \alpha_2 \\ \vdots \\ \cos \alpha_n \end{pmatrix} = \frac{\mathfrak{x}}{|\mathfrak{x}|} = \frac{\mathfrak{x}}{R} \tag{4}$$

ein. Man erkennt, daß (1) die konvexe Hülle der Punkte $\pm \mathfrak{r}_i$ ist. Der Ausdruck

$$J = \sum_{i=1}^{p} (\mathfrak{r}_i^T \mathfrak{c})^2 \tag{5}$$

läßt sich als „Trägheitsmoment" bezüglich der zur Richtung $\mathfrak{c}$ orthogonalen Hyperebene $\mathfrak{c}^T \mathfrak{x} = 0$ bezeichnen. Es ist

$$\begin{aligned} J &= \sum_{i=1}^{p} \left(\sum_{k=1}^{n} b_{ik} \cos \alpha_k \right)^2 \\ &= \sum_{i=1}^{p} \sum_{k,l=1}^{n} b_{ik} b_{il} \cos \alpha_k \cos \alpha_l \\ &= \sum_{k,l=1}^{n} \left(\sum_{i=1}^{p} b_{ik} b_{il} \right) \cos \alpha_k \cos \alpha_l . \end{aligned}$$

Damit geht (5) über in

$$J = \sum_{k,l=1}^{n} J_{kl} \cos\alpha_k \cos\alpha_l, \tag{6}$$

wo

$$J_{kl} = \sum_{i=1}^{p} b_{ik} b_{il} \quad (k, l = 1, \ldots, n) \tag{7}$$

gesetzt ist. Trägt man nun in jeder Richtung $\mathfrak{c}$ den Vektor der Länge

$$R = \frac{1}{\sqrt{J}} \tag{8}$$

ab, wo J das zur Richtung $\mathfrak{c}$ gehörige Trägheitsmoment ist, so genügen die Endpunkte dieser Vektoren der Beziehung

$$J = \frac{1}{R^2} = \sum_{k,l=1}^{n} J_{kl} \frac{x_k x_l}{R^2}, \tag{9}$$

d.h. sie erfüllen die Hyperfläche zweiter Ordnung

$$\sum_{k,l=1}^{n} J_{kl} x_k x_l = 1, \quad J_{kl} = \sum_{i=1}^{p} b_{ik} b_{il}. \tag{10}$$

Diese Fläche läßt sich durch eine orthogonale Substitution der x_i stets auf Hauptachsen transformieren, wobei die Punkte $\pm\mathfrak{r}_i$ entsprechend mittransformiert werden. Wir haben also erhalten:

Satz 6.1. *Jedes konvexe Mittelpunktspolytop läßt sich durch eine Gleichung*

$$\max_{i=1}^{p} \left| \sum_{k=1}^{n} b_{ik} u_k \right| = 1 \quad \text{mit} \quad \sum_{i=1}^{p} b_{ik} b_{il} = 0 \quad \text{für} \quad k \neq l \tag{11}$$

darstellen.

Durch Dualisierung erhält man:

Satz 6.2. *Jedes konvexe Mittelpunktspolytop läßt sich durch eine Gleichung*

$$\max_{i=1}^{m} \left| \sum_{k=1}^{n} a_{ik} x_k \right| = 1 \quad \text{mit} \quad \sum_{i=1}^{m} a_{ik} a_{il} = 0 \quad \text{für} \quad k \neq l \tag{12}$$

darstellen.

Man erkennt unmittelbar, daß man zu einem beliebigen konvexen Mittelpunktspolytop mit $2p$ Ecken weitere Paare symmetrisch zum Nullpunkt gelegener, im Inneren oder teilweise auf dem Rande befindlicher Punkte derart hinzunehmen kann, daß aus dem Trägheits-„ellipsoid" eine „Trägheitshyperkugel" entsteht. Dies bedeutet nach Satz 4.2, daß jedes konvexe Mittelpunktspolytop als senkrechte Parallelprojektion eines höherdimensionalen Maßpolytops darstellbar ist. Wir führen nun eine entsprechende Rechnung mit dem Ziel einer Abschätzung durch, und zwar zunächst für den dualen Fall.

Sei ein konvexes Mittelpunktspolytop gegeben, und zwar bereits in der „auf Hauptachsen transformierten" Form (12). Es möge das Maßpolytop $\sum_{i=1}^{n} |x_i| = r$ einschließen und von dem Würfel $\max_{i=1}^{n} |x_i| = R$ eingeschlossen werden. Wir betrachten die Koeffizientenmatrix $((a_{ik}))$. Nach Voraussetzung verschwinden die Summen $\sum_{i=1}^{m} a_{ik} a_{il}$ für $k \neq l$. Zur Abkürzung setzen wir $\sum_{i=1}^{m} a_{ik}^2 = A_k$ und nehmen ohne Beschränkung der Allgemeinheit an, daß $A_1 \geq A_k$ für $k = 1, \ldots, n$ ist. Nach Definition von r folgt $|a_{ik}| \leq \frac{1}{r}$ für $i = 1, \ldots, m$, $k = 1, \ldots, n$, also $A_1 \leq \frac{m}{r^2}$. Sei nun

$$t = \left[m \cdot \frac{R^2}{r^2} \right], \tag{13}$$

d.h. die kleinste natürliche Zahl t mit $t \geq m \cdot \frac{R^2}{r^2}$. Dann ist

$$\max_{k=1}^{n} \left| \sqrt{\frac{A_1 - A_k}{t}} \cdot x_k \right| < \frac{\sqrt{\frac{m}{r^2}}}{\sqrt{\frac{m R^2}{r^2}}} \cdot R = 1 \tag{14}$$

für alle Punkte des gegebenen Polytops, folglich ist das Polytop

$$\max_{i=1}^{m+t(n-1)} \left| \sum_{k=1}^{n} a'_{ik} x_k \right| = 1 \tag{15}$$

mit

$$a'_{ik} = \begin{cases} a_{ik} & \text{für} \quad i = 1, \ldots, m, \; k = 1, \ldots, n \\ \sqrt{\frac{A_1 - A_k}{t}} & \text{für} \quad \begin{cases} k = 2, \ldots, n \\ i = m + (k-2)\,t + 1, \ldots, m + (k-1)\,t, \end{cases} \\ 0 & \text{in allen übrigen Fällen} \end{cases}$$

mit dem gegebenen Polytop identisch, und es ist

$$\sum_{i=1}^{m+t(n-1)} a'_{ik} a'_{il} = 0 \quad \text{für} \quad k \neq l, \; k, l = 1, \ldots, n,$$

$$\sum_{i=1}^{m+t(n-1)} a'^2_{i1} = \sum_{i=1}^{m} a_{i1}^2 = A_1,$$

$$\sum_{i=1}^{m+t(n-1)} a'^2_{ik} = \sum_{i=1}^{m} a_{ik}^2 + t \cdot \frac{A_1 - A_k}{t} = A_1 \quad \text{für} \quad k = 2, \ldots, n.$$

Folglich haben wir nach Satz 4.1 erhalten:

Satz 6.3. *Jedes konvexe Mittelpunktspolytop des R_n mit $2m$ extremen Stützen, das bei einer bestimmten Lage im Koordinatensystem das Maßpolytop $\sum_{i=1}^{n} |x_i| = r$ einschließt und von dem Würfel $\max_{i=1}^{n} |x_i| = R$ eingeschlossen wird,*

ist als Mittelpunktsschnitt des $(m+m\alpha^2\cdot(n-1))$-dimensionalen Würfels darstellbar, wo $\alpha \geq \frac{R}{r}$ und $m\alpha^2$ eine natürliche Zahl ist.

Durch Dualisierung folgt hieraus:

Satz 6.4. *Jedes konvexe Mittelpunktspolytop des R_n mit $2p$ Ecken, das bei einer bestimmten Lage im Koordinatensystem das Maßpolytop $\max_{i=1}^{n} |u_i| = R$ einschließt und von dem Würfel $\sum_{i=1}^{n} |u_i| = r$ eingeschlossen wird, ist als senkrechte Parallelprojektion des $(p+p\alpha^2\cdot(n-1))$-dimensionalen Maßpolytops darstellbar, wo $\alpha \geq \frac{R}{r}$ und $p\alpha^2$ eine natürliche Zahl ist.*

Auf dem in diesem Paragraphen beschrittenen Wege ließen sich auch entsprechende Abschätzungen für beliebige konvexe Polytope herleiten, da man z.B. jedes konvexe Polytop des R_n in einfacher Weise als Schnitt eines $(n+1)$-dimensionalen Mittelpunktspolytops darstellen kann. Wir wollen an dieser Stelle aber nicht näher darauf eingehen.

7. In diesem Paragraphen wird eine weitere analytische Darstellung konvexer Polytope gegeben.

Seien

$$\sum_{k=1}^{n} a_{ik} x_k - 1 = 0 \qquad (i = 1, \ldots, m) \tag{1}$$

die extremen Stützen eines konvexen Polytops; der Koordinatenursprung sei innerer Punkt des Polytops. Gesucht sind positive „Gewichte" $p_1, p_2, \ldots, p_m$, so daß

$$\sum_{i=1}^{m} p_i \left(\sum_{k=1}^{n} a_{ik} x_k - 1 \right) \equiv \text{const} \tag{2}$$

ist für den gesamten n-dimensionalen Raum. Durch Umordnung dieses Ausdrucks ergibt sich

$$\sum_{k=1}^{n} \left(\sum_{i=1}^{m} p_i a_{ik} \right) x_k - \sum_{i=1}^{m} p_i \equiv \text{const}, \tag{3}$$

also

$$\sum_{i=1}^{m} p_i a_{ik} = 0 \qquad (k = 1, \ldots, n) \tag{4}$$

als Bedingungen für die p_i. Nun sind

$$\sum_{k=1}^{n} a_{ik} u_k = 1, \tag{5}$$

d.h. die Punkte mit den Koordinaten

$$(a_{i1}, a_{i2}, \ldots, a_{in}) \quad \text{mit} \quad i = 1, \ldots, m \tag{6}$$

die Ecken des zum ursprünglichen Polytop polaren Polytops, es ist wie das ursprüngliche konvex und besitzt den Koordinatenursprung als inneren

Punkt. Es ist daher der Punkt $(0, 0, \ldots, 0)$ als Schwerpunkt einer positiven Massenbelegung der Punkte (6) darstellbar, d.h. das Gleichungssystem (4) und damit auch die Beziehung (2) besitzen ein System positiver Lösungen p_i.

Nachdem die Existenz der positiven Lösungen p_i gezeigt ist, zeigen wir:

Satz 7.1. *Sind* $\sum_{k=1}^{n} a_{ik} x_k = 1$ $(i = 1, \ldots, m)$ *die extremen Stützen eines konvexen Polytops, das den Punkt* $(0, 0, \ldots, 0)$ *als inneren Punkt enthält, so lassen sich stets (im allgemeinen auf mehrfache Weise) positive Gewichte* p_i *finden, daß*

$$\frac{\sum_{i=1}^{m} \left| p_i \left(\sum_{k=1}^{n} a_{ik} x_k - 1 \right) \right|}{\sum_{i=1}^{m} p_i} = 1 \tag{7}$$

eine Darstellung der Randpunkte und inneren Punkte des gegebenen Polytops ist.

Beweis. Für Rand- und innere Punkte sind die linken Seiten der Gln. (1) nicht positiv, also ist, wenn man für die p_i ein System der oben gefundenen Gewichte einsetzt,

$$\sum_{i=1}^{m} \left| p_i \left(\sum_{k=1}^{n} a_{ik} x_k - 1 \right) \right| = - \sum_{i=1}^{m} p_i \left(\sum_{k=1}^{n} a_{ik} x_k - 1 \right) = \sum_{i=1}^{m} p_i \tag{8}$$

und die Gl. (7) folglich erfüllt. Für jeden äußeren Punkt gilt nach passender Umnumerierung der extremen Stützen

$$\sum_{k=1}^{n} a_{ik} x_k - 1 \begin{cases} > 0 \text{ für } i = 1, \ldots, j \quad (j \geq 1) \\ \leq 0 \text{ für } i = j+1, \ldots, m, \end{cases} \tag{9}$$

also

$$\begin{cases} \sum_{i=1}^{m} \left| p_i \left(\sum_{k=1}^{n} a_{ik} x_k - 1 \right) \right| = \sum_{i=1}^{j} p_i \left(\sum_{k=1}^{n} a_{ik} x_k - 1 \right) - \sum_{i=j+1}^{m} p_i \left(\sum_{k=1}^{n} a_{ik} x_k - 1 \right) \\ \qquad = - \sum_{i=1}^{m} p_i \left(\sum_{k=1}^{n} a_{ik} x_k - 1 \right) + 2 \sum_{i=1}^{j} p_i \left(\sum_{k=1}^{n} a_{ik} a_k - 1 \right) \\ \qquad > \sum_{i=1}^{m} p_i, \end{cases} \tag{10}$$

womit die Behauptung bewiesen ist.

Nach dem Dualitätsprinzip ergibt sich:

Satz 7.2. *Sind* $(b_{i1}, b_{i2}, \ldots, b_{in})$ *mit* $i = 1, \ldots, p$ *die Ecken eines konvexen Polytops, das den Punkt* $(0, 0, \ldots, 0)$ *als inneren Punkt enthält, so lassen sich stets positive Gewichte* p_i *finden, daß*

$$\frac{\sum_{i=1}^{p} \left| p_i \left(\sum_{k=1}^{n} b_{ik} u_k - 1 \right) \right|}{\sum_{i=1}^{p} p_i} = 1 \tag{11}$$

eine Darstellung der einhüllenden und meidenden Hyperebenen des gegebenen Polytops ist.

8. Die im letzten Paragraphen gewonnene Darstellung (7.7) werten wir in einem speziellen Falle weiter aus. Der n-dimensionale Würfel der Kantenlänge 2 läßt sich nach diesem Formalismus in der Gestalt

$$(1)\quad |1+x_1|+|1-x_1|+|1+x_2|+|1-x_2|+\cdots+|1+x_n|+|1-x_n|=2n$$

darstellen. Die Beziehung (1) läßt sich aber auch unmittelbar einsehen. Es zeigt sich nun, daß dieser Würfel als Schnitt des $2n$-dimensionalen Maßpolytops

$$(2)\qquad |y_1|+|y_2|+\cdots+|y_{2n}|=n\cdot\sqrt{2}$$

erhalten werden kann, und zwar als Schnitt eines der dieses Maßpolytop berandenden $(2n-1)$-dimensionalen Simplizes der Kantenlänge $2n$. Denn wendet man auf (2) die orthogonale Substitution

$$(3)\qquad y_i=\begin{cases}\dfrac{x_i}{\sqrt{2}}+\dfrac{x_{n+i}}{\sqrt{2}} & \text{für}\quad i=1,\ldots,n\\[2ex] -\dfrac{x_{i-n}}{\sqrt{2}}+\dfrac{x_i}{\sqrt{2}} & \text{für}\quad i=n+1,\ldots,2n\end{cases}$$

an und schneidet man das so transformierte Maßpolytop mit dem n-dimensionalen linearen Raum

$$x_{n+1}=x_{n+2}=\cdots=x_{2n}=1,$$

so ergibt sich in der Tat der Würfel (1). Da der Gl. (1) außer den Randpunkten auch alle inneren Punkte des Würfels genügen, der Gl. (2) aber lediglich die Randpunkte des Maßpolytops, so folgt, daß (1) durch Schnitt eines Randpolytops von (2), also eines regulären $(2n-1)$-dimensionalen Simplex der Kantenlänge $2n$, dargestellt ist.

Der Schwerpunkt des Simplex mit den Ecken

$$\begin{aligned}\mathfrak{y}_1&=\left(\frac{2n}{\sqrt{2}},\,0,\ldots,0\right)\\ \mathfrak{y}_2&=\left(0,\,\frac{2n}{\sqrt{2}},\ldots,0\right)\\ &\vdots\qquad\qquad\ddots\\ \mathfrak{y}_{2n}&=\left(0,0,\ldots,\frac{2n}{\sqrt{2}}\right),\end{aligned}$$

also der Punkt $\mathfrak{y}=\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},\ldots,\frac{1}{\sqrt{2}}\right)$, geht hierbei in den Mittelpunkt des Würfels über.

Es gilt also:

Satz 8.1. *Der n-dimensionale Würfel der Kantenlänge a läßt sich darstellen als Schnitt durch den Schwerpunkt eines $(2n-1)$-dimensionalen regulären Simplex der Kantenlänge $n\cdot a$ mit einem n-dimensionalen linearen Raum.*

Der hierzu duale Satz ist von P. H. SCHOUTE [2] § 8.84 (Theorem II) bewiesen worden.

Nach Satz 4.3 folgt:

Satz 8.2. *Jedes konvexe Hauptpolytop, das von $2m$ extremen Stützen begrenzt wird und $2p$ Ecken besitzt, ist als Schnitt durch den Schwerpunkt eines $(2m-1)$-dimensionalen regulären Simplex und als senkrechte Parallelprojektion eines $(2p-1)$-dimensionalen regulären Simplex darstellbar.*

Aus Satz 5.1 und Satz 5.2 ergibt sich:

Satz 8.3. *Jedes konvexe Polytop ist als Schnitt eines höherdimensionalen regulären Simplex und als Projektion eines höherdimensionalen regulären Simplex darstellbar.*

Aus Satz 6.3 und Satz 6.4 ergibt sich:

Satz 8.4. *Jedes konvexe Mittelpunktspolytop ist als Schnitt durch den Schwerpunkt eines höherdimensionalen regulären Simplex und als senkrechte Parallelprojektion eines höherdimensionalen regulären Simplex darstellbar.*

Satz 8.3 und Satz 8.4 lassen sich leicht durch Angabe der entsprechenden Abschätzungen ergänzen.

Literatur

[1] Bonnesen, T., u. W. Fenchel: Theorie der konvexen Körper. Berlin 1934. — [2] Schoute, P. H.: Mehrdimensionale Geometrie, Bd. II. Leipzig 1905.

Schwelm (Westf.), Tannenstr. 8

(Eingegangen am 22. November 1955)

LANDSBERG, M.
Math. Zeitschr. Bd. 65, S. 104—112 (1956)

Lineare topologische Räume, die nicht lokalkonvex sind

Von
MAX LANDSBERG

Einleitung und Ergebnisse

Bekanntlich sind für die Funktionalanalysis die lokalkonvexen Räume von fundamentaler Bedeutung. In den beiden letzten Jahrzehnten hat man jedoch auch spezielle lineare topologische Räume, die nicht lokalkonvex sind, näher untersucht (s. z.B. M. M. DAY [2], S. S. WALTERS [9]). Der Nachweis, daß es sich hierbei um nicht lokalkonvexe Räume handelt, war nicht in allen Fällen einfach. So behandelte S. S. WALTERS eingehend die Räume H' für $0<r<1$ ([9], [10]), konnte aber seine Vermutung, daß diese Räume nicht lokalkonvex sind, nicht beweisen (s. [10]). Die Richtigkeit der Annahme von WALTERS wies A. E. LIVINGSTON [4] nach.

Die vorliegende Note enthält unter anderem ein einfaches Kriterium, mit dem sich in vielen Fällen die Tatsache, daß ein vorliegender linearer topologischer Raum nicht lokalkonvex ist, leicht feststellen läßt.

Der Ausgangspunkt für die hier durchgeführten Überlegungen ist die folgende Definition. Eine Teilmenge M eines linearen Raumes soll r-konvex $(0<r\leq 1)$ heißen, wenn aus $a\geq 0$, $b\geq 0$, $a^r+b^r=1$, $x\in M$, $y\in M$ stets $ax+by\in M$ folgt. Wir bezeichnen einen linearen topologischen Raum als r-konvex, wenn jede Nullumgebung eine r-konvexe Nullumgebung enthält. Offensichtlich ist die Klasse der 1-konvexen Räume mit der Gesamtheit der lokalkonvexen Räume identisch. Die meisten linearen topologischen Räume, die bisher zu funktionalanalytisch interessanten Aussagen Anlaß gaben, sind für ein gewisses r $(0<r\leq 1)$ r-konvex.

Man erkennt leicht, daß jeder r-konvexe Raum auch s-konvex ist für jedes s mit $0<s<r$. Dagegen ist ein r-konvexer Raum mit $0<r<1$ im allgemeinen nicht s-konvex für $r<s\leq 1$. Bezüglich dieser Bemerkung gilt nun der folgende Sachverhalt: *Es sei E ein linearer topologischer Raum und φ ein im Nullelement von E stetiges Funktional (auf E). Außerdem existiere ein r mit $0<r<1$ und eine beschränkte Folge (v_n) aus E, so daß für jedes natürliche n und für beliebige positive Zahlen $\varrho_1, \ldots, \varrho_n$*

$$\varphi(\varrho_1 v_1+\cdots+\varrho_n v_n)\geq \varrho_1^r+\cdots+\varrho_n^r$$

ist. Dann ist E nicht s-konvex für $r<s\leq 1$, insbesondere also nicht lokalkonvex.

An verschiedenen Beispielen soll die Anwendbarkeit dieses Kriteriums dargelegt werden. Im Anschluß an [4], [9] und [10] behandeln wir zunächst einen relativ einfachen Fall. Es sei $0<r<1$ und G ein linearer Funktionenraum, dessen Elemente komplexwertige Funktionen sind, die im Innern des

Einheitskreises der komplexen z-Ebene definiert und dort stetig sind. Insbesondere gehöre zu G jede Funktion, die für $|z|<1$ mit einem Polynom $a_0+a_1z+\cdots+a_mz^m$ übereinstimmt. Alle $f\in G$ mit

$$q(f)=\sup_{t\in(0,1)}\int_0^{2\pi}|f(te^{i\vartheta})|^r\,d\vartheta<+\infty$$

bilden dann einen linearen Funktionenraum $H^r(G)$, den die Metrik $q(f-g)$ $(f, g\in H^r(G))$ zu einem linearen topologischen Raum macht. Es wird gezeigt, daß $H^r(G)$ für kein s mit $r<s\leqq 1$ s-konvex sein kann.

Im nächsten Fall sei F ein linearer Folgenraum, der alle finiten Folgen enthalten soll, und (p_n) eine Folge von reellen Zahlen mit $0<p_n\leqq 1$ $(n=1, 2, \ldots)$, $0<\liminf_{n\to\infty} p_n<1$. Alle $x=(a_n)$ aus F mit

$$p(x)=\sum_{n=1}^{\infty}|a_n|^{p_n}<+\infty$$

bilden dann einen linearen Folgenraum $l(F;(p_n))$, der durch die Metrik $p(x-y)$ $(x, y\in l(F;(p_n)))$ zu einem linearen topologischen Raum wird. Es zeigt sich, daß $l(F;(p_n))$ r-konvex für jedes r mit $0<r<\liminf_{n\to\infty} p_n$ ist, aber für kein s mit $\liminf_{n\to\infty} p_n<s\leqq 1$ s-konvex sein kann.

Für $F=\omega$, wo ω der Raum aller Folgen ist, findet man einige Angaben über $l(\omega;(p_n))=l((p_n))$ bei H. NAKANO [6].

Für $F=\omega$ und $p_n=r$ $(0<r<1;\ n=1, 2, \ldots)$ haben wir es mit dem Raum l_r zu tun. Daß l_r nicht lokalkonvex ist, ist bekannt (s. z.B. MAZUR-ORLICZ [5], A. TYCHONOFF [8] $(r=1/2)$). Man kann l_r noch unter einem anderen Gesichtspunkt betrachten. Es sei Z die Menge aller natürlichen Zahlen, S_0 der σ-Ring aller Teilmengen von Z und μ_0 dasjenige Maß auf S_0, das durch $\mu_0(\{n\})=1$ $(n=1, 2, \ldots)$ bestimmt ist. Dann ist l_r mit dem zu Z, S_0, μ_0 gehörigen Raum L_r identisch und für jedes $f\in L_r$ ist

$$\int|f|^r\,d\mu_0=\sum_{n=1}^{\infty}|f(n)|^r.$$

Im Spezialfall Z, S_0, μ_0 ist also L_r $(0<r<1)$ für kein s mit $r<s\leqq 1$ s-konvex. Daß in gewissen Sonderfällen L_r $(0<r<1)$ nicht lokalkonvex ist, ist bekannt (s. z.B. M. M. DAY [2], MAZUR-ORLICZ [5]). Ist allgemeiner S ein σ-Ring auf einer Menge E und μ ein Maß auf S, so wird man fragen, unter welchen weiteren Bedingungen für E, S, μ die zugehörigen Räume L_r $(0<r<1)$ nicht s-konvex für $r<s\leqq 1$ sind. Mit Hilfe des oben angeführten Satzes ergibt sich:

1. E sei ein lokalkompakter Raum mit unendlich vielen Elementen, der σ-Ring S enthalte das System K_0 aller kompakten Teilmengen von E, die ein G_δ sind. Für jedes $A\in K_0$, das einen inneren Punkt besitzt, sei $\mu(A)$ endlich und positiv. Dann ist L_r $(0<r<1)$ für kein s mit $r<s\leqq 1$ s-konvex.

2. E sei eine unendliche (bezüglich S, μ) meßbare Gruppe. Auch dann kann L_r $(0<r<1)$ für kein $s\in(r, 1]$ s-konvex sein.

Im zweiten Fall, wo keine Topologie auf E gegeben ist, erscheint es zunächst problematisch, ob man das oben angegebene Kriterium anwenden kann. Die Einführung der WEILschen Topologie, durch die E zu einer topologischen Gruppe mit einer lokalkompakten vollständigen Hülle wird, beseitigt jedoch jede Schwierigkeit.

Lineare topologische Räume, die r-konvex sind

Wir betrachten nur lineare Räume, die komplex sind. Der Begriff Topologie ist im Sinne von BOURBAKI (N. BOURBAKI [1]) zu verstehen, jedoch soll (im Gegensatz zu BOURBAKI) ein topologischer Raum immer ein HAUSDORFF-Raum sein.

In einem linearen Raum wollen wir eine Teilmenge als absolut r-konvex bezeichnen, wenn sie r-konvex (s. Einleitung) ist und mit x auch immer $e^{i\varrho}x$ (ϱ reell) enthält.

Bemerkung 1. *Der lineare topologische Raum E sei r-konvex und U eine Nullumgebung. Dann enthält U eine Nullumgebung, die offen und absolut r-konvex ist.*

Beweis. U enthält eine abgeschlossene und r-konvexe Nullumgebung V. V wiederum enthält eine Nullumgebung W, zu der mit x stets auch $e^{i\varrho}x$ (ϱ reell) gehört. Die r-konvexe Hülle von W, d.h. die kleinste W enthaltende r-konvexe Teilmenge von E, sei W_0. Berücksichtigt man $\overline{W}_0 \subset V \subset U$, $\overline{W}_0$ ist die abgeschlossene Hülle von W_0, so sieht man, daß das Innere von $\overline{W}_0$ eine offene und absolut r-konvexe Nullumgebung ist, die ganz in U liegt.

Ein Funktional auf dem linearen Raum E ist eine Abbildung von E in die Menge der reellen Zahlen. Das Funktional φ wird als positiv-homogen bzw. als subadditiv bezeichnet, wenn $\varphi(tx) = t\varphi(x)$ $(x \in E,\ t > 0)$ bzw. $\varphi(x+y) \leq \varphi(x) + \varphi(y)$ $(x, y \in E)$ ist. Das Funktional φ heiße eine r-Halbnorm, wenn φ subadditiv und $\varphi(ax) = |a|^r \varphi(x)$ ($x \in E$, a komplex) ist. Ist außerdem $\varphi(x) \neq 0$ für jedes vom Nullelement verschiedene x aus E, so sprechen wir von einer r-Norm.

Bemerkung 2. *E sei ein linearer topologischer Raum und U eine offene und absolut r-konvexe Nullumgebung. Dann existiert genau eine r-Halbnorm φ auf E mit*

$$U = \{x \in E : \varphi(x) < 1\}.$$

Beweis. In naheliegender Weise setzen wir

$$\varphi(x) = \inf\{\varrho^r : \varrho > 0 \quad \text{und} \quad (1/\varrho)\, x \in U\}.$$

Mit einfacher Rechnung bestätigt man $\varphi(ax) = |a|^r \varphi(x)$ ($x \in E$, a komplex) und $U = \{x \in E : \varphi(x) < 1\}$. Weiter gibt es zu jedem $x \in E$ bzw. $y \in E$ ein $\varrho > 0$ bzw. $\sigma > 0$ mit $(1/\varrho)\, x \in U$ bzw. $(1/\sigma)\, y \in U$. Dann gehört auch

$$\frac{\varrho}{(\varrho^r + \sigma^r)^{1/r}} \cdot \frac{1}{\varrho} x + \frac{\sigma}{(\varrho^r + \sigma^r)^{1/r}} \cdot \frac{1}{\sigma} y = \frac{1}{(\varrho^r + \sigma^r)^{1/r}} (x + y)$$

zur (r-konvexen) Menge U. Also ist $\varphi(x+y) < \varrho^r + \sigma^r$ z.B. auch für $\varrho = (\varphi(x)+1/n)^{1/r}$, $\sigma = (\varphi(y)+1/n)^{1/r}$ bei beliebigem natürlichem n. Hieraus folgt unmittelbar die Subadditivität von φ. Die Eindeutigkeitsaussage ergibt sich sofort aus dem folgenden Hilfssatz, dessen Beweis wir uns schenken können.

Hilfssatz 1. *φ und ψ seien zwei positiv-homogene Funktionale auf dem linearen Raum E. Außerdem sei $\varphi(x) \geq 0$ $(x \in E)$ und aus $\varphi(x) < 1$ folge stets $\psi(x) < 1$. Dann ist $\psi(x) \leq \varphi(x)$ für alle x aus E.*

Bemerkung 3. *In einem linearen Raum sei M eine r-konvexe Teilmenge, die mit x auch immer ϱx $(0 < \varrho < 1)$ enthält. Dann ist M auch s-konvex für $0 < s < r$.*

Beweis. Es sei $0 < s < r$, $a \geq 0$, $b \geq 0$, $a^s + b^s = 1$, $x \in M$, $y \in M$. Dann gehört auch

$$y_0 = (a/q)\, x + (b/q)\, y$$

zu M für $q = (a^r + b^r)^{1/r}$. Denn es ist $(a/q)^r + (b/q)^r = 1$. Nun ist

$$q^s = (a^r + b^r)^{s/r} \leq a^s + b^s = 1,$$

also $q \leq 1$. Nach Voraussetzung gehört damit auch $q\, y_0 = a\, x + b\, y$ zu M, M ist s-konvex.

Aus dieser Bemerkung folgt sofort

Satz 1. *E sei ein r-konvexer Raum. Dann ist E auch s-konvex für jedes s mit $0 < s < r$. Insbesondere ist also jeder lokalkonvexe Raum auch s-konvex für $0 < s < 1$.*

Wir erinnern daran, daß in einem linearen topologischen Raum eine Teilmenge A als beschränkt bezeichnet wird, wenn es zu jeder Nullumgebung U ein gewisses $\varrho > 0$ mit $A \subset \varrho U$ gibt.

Satz 2. *E sei ein linearer topologischer Raum, φ ein im Nullelement von E stetiges Funktional. Außerdem existiere eine beschränkte Folge (v_n) aus E und ein r mit $0 < r < 1$, so daß für jede natürliche Zahl n und für beliebige positive Zahlen $\varrho_1, \ldots, \varrho_n$ stets*

$$\varphi(\varrho_1 v_1 + \cdots + \varrho_n v_n) \geq \varrho_1^r + \cdots + \varrho_n^r$$

ist. Dann kann E nicht s-konvex für $r < s \leq 1$ sein. Insbesondere ist E nicht lokalkonvex.

Beweis. E sei doch s-konvex für ein gewisses s mit $r < s \leq 1$. Nun ist $U = \{x \in E : \varphi(x) < 1 + \varphi(o)\}$ eine Umgebung vom Nullelement o des Raumes E. Nach Bemerkung 1 enthält U eine Nullumgebung V, die offen und absolut s-konvex ist. Die mit V nach Bemerkung 2 assoziierte s-Halbnorm sei Φ. Nach Voraussetzung existiert ein $k > 0$, so daß $\Phi(v_n) \leq k$ für alle n ist. Es sei

$$w_n = v_1 + (1/2^{1/r})\, v_2 + \cdots + (1/n^{1/r})\, v_n \qquad (n = 1, 2, \ldots).$$

Dann ist

$$\begin{aligned} \Phi(w_n) &\leq \Phi(v_1) + (1/2^{s/r})\, \Phi(v_2) + \cdots + (1/n^{s/r})\, \Phi(v_n) \\ &\leq k\,(1 + 1/2^{s/r} + \cdots + 1/n^{s/r}), \end{aligned}$$

also $\Phi(w_n) < k\,\zeta\left(\frac{s}{r}\right)$ für jedes n. Das Element

$$w_n' = \left(\frac{1}{k\,\zeta\left(\frac{s}{r}\right)}\right)^{1/s} w_n$$

gehört somit für jedes n zu V, also auch zu U. Das ist aber nicht möglich, da wegen

$$\varphi(w_n') \geq \left(\frac{1}{k\,\zeta\left(\frac{s}{r}\right)}\right)^{r/s} (1 + 1/2 + \cdots + 1/n)$$

nicht für jedes n $\varphi(w_n') < 1 + \varphi(o)$ sein kann.

Die Räume $H^r(G)$ $(0 < r < 1)$

Satz 3. *$H^r(G)$ $(0 < r < 1$; s. Einleitung$)$ ist nicht s-konvex für $r < s \leq 1$. Insbesondere ist also $H^r(G)$ nicht lokalkonvex.*

Beweis (vgl. A. E. LIVINGSTON [4]). C sei der lineare Raum aller auf $0 \leq \vartheta \leq 2\pi$ erklärten und dort stetigen komplexwertigen Funktionen. Die Norm $\|h\| = \sup_{\vartheta \in [0, 2\pi]} |h(\vartheta)|$ $(h \in C)$ macht C zu einem BANACH-Raum, durch

$$\hat{q}(h) = \int_0^{2\pi} |h(\vartheta)|^r \, d\vartheta$$

ist auf C eine wegen

(1) $$\hat{q}(h) \leq 2\pi \|h\|^r \qquad (h \in C)$$

stetige r-Norm (s. Einleitung) gegeben. Zu jedem natürlichen n gibt es ein $h_n \in C$ mit $\hat{q}(h_n) = 2$, das auf den Intervallen $\left[0, 2\pi \frac{n}{n+1}\right]$, $\left[2\pi \frac{n+1}{n+2}, 2\pi\right]$ verschwindet. Offensichtlich gilt dann für jedes n und für beliebige komplexe Zahlen $a_1, \ldots, a_n$

(2) $$\hat{q}(a_1 h_1 + \cdots + a_n h_n) = 2(|a_1|^r + \cdots + |a_n|^r).$$

Weiter existiert nach einem bekannten Satz zu jedem h_n ein trigonometrisches Polynom

$$t_n(\vartheta) = \sum_{k=-s_n}^{s_n} a_{kn} e^{ik\vartheta}$$

mit

(3) $$\|t_n - h_n\| \leq \left(\frac{1}{2\pi}\right)^{1/r}.$$

Mit Hilfe von t_n bilden wir das Polynom

$$P_n(z) = \sum_{k=-s_n}^{s_n} a_{kn} z^{k+s_n},$$

dessen Einschränkung auf $|z| < 1$ mit π_n bezeichnet werden soll. Nach Voraussetzung gehört π_n zu G (s. Einleitung) und daher auch zu $H^r(G)$. Berücksichtigt man (1), (3) und einen bekannten funktionentheoretischen Satz

(PÓLYA-SZEGÖ [7], S. 144), so findet man für jedes n

$$q(\pi_n) = \sup_{t\in(0,1)} \int_0^{2\pi} |P_n(t e^{i\vartheta})|^r \, d\vartheta = \lim_{t\to 1} \int_0^{2\pi} |P_n(t e^{i\vartheta})|^r \, d\vartheta$$

$$= \hat{q}(t_n) \leq 2\pi \|t_n - h_n\|^r + \hat{q}(h_n) \leq 3.$$

Die Folge (π_n) ist also beschränkt im Raum $H^r(G)$. Wir zeigen noch, daß für jedes n und für beliebige positive Zahlen $\varrho_1, \ldots, \varrho_n$

$$q(\varrho_1\pi_1 + \cdots + \varrho_n\pi_n) \geq \varrho_1^r + \cdots + \varrho_n^r$$

ist, womit nach Satz 2 alles bewiesen ist. Setzen wir $g_n(\vartheta) = e^{i s_n \vartheta}$ $(n = 1, 2, \ldots;$ $0 \leq \vartheta \leq 2\pi)$, so ist mit Berücksichtigung von (1), (2), (3)

$$q(\varrho_1\pi_1 + \cdots + \varrho_n\pi_n) = \hat{q}\left(\sum_{j=1}^{n} \varrho_j g_j t_j\right)$$

$$\geq \hat{q}\left(\sum_{j=1}^{n} \varrho_j g_j h_j\right) - \hat{q}\left(\sum_{j=1}^{n} \varrho_j g_j (t_j - h_j)\right)$$

$$\geq 2\sum_{j=1}^{n} \varrho_j^r - 2\pi \sum_{j=1}^{n} \varrho_j^r \|t_j - h_j\|^r$$

$$\geq \sum_{j=1}^{n} \varrho_j^r.$$

Die Räume $l(F; (p_n))$

Im Anschluß an die in der Einleitung angegebenen Bezeichnungen setzen wir noch $\lambda = \liminf_{n\to\infty} p_n$.

Satz 4. *$l(F; (p_n))$ ist r-konvex für jedes r mit $0 < r < \lambda$.*

Beweis. Es sei r eine Zahl aus $(0, \lambda)$. Es gibt dann eine natürliche Zahl n_0 mit $p_n \geq r$ für alle $n > n_0$. Für $x = (a_n)$ aus $l(F; (p_n))$ sei

$$\hat{p}(x) = |a_1|^r + \cdots + |a_{n_0}|^r + \sum_{n=n_0+1}^{\infty} |a_n|^{p_n}.$$

$\hat{p}$ ist ein subadditives Funktional auf $l(F; (p_n))$, für das insbesondere $\hat{p}(a x) \leq |a|^r \hat{p}(x)$ (a komplex, $|a| \leq 1$, $x \in l(F; (p_n))$) gilt. Für jedes natürliche m sei $U_m = \{x \in l(F; (p_n)) : \hat{p}(x) \leq 1/m\}$. Die U_m bilden eine Basis des Filters aller Nullumgebungen. Wir zeigen noch, daß jedes U_m r-konvex ist. Es sei $a, b \geq 0$ und $a^r + b^r = 1$. Für zwei beliebige Elemente x, y aus U_m ist dann

$$\hat{p}(a x + b y) \leq a^r \hat{p}(x) + b^r \hat{p}(y) \leq (a^r + b^r) \cdot \frac{1}{m} = \frac{1}{m}.$$

Also gehört auch $a x + b y$ zu U_m.

Satz 5. *$l(F; (p_n))$ ist nicht s-konvex für $\lambda < s \leq 1$. Insbesondere ist $l(F; (p_n))$ nicht lokalkonvex.*

Beweis. s gehöre zum Intervall $(\lambda, 1]$. Es wird ein $\varkappa$ mit $\lambda<\varkappa<s$ gewählt. Es gibt dann eine Teilfolge (p_{n_j}) von (p_n) mit $p_{n_j}\leq\varkappa$ $(j=1, 2, \ldots)$. Für jedes $x=(a_n)$ aus $l(F; (p_n))$ ist

$$\tilde{p}(x)=\sum_{j=1}^{\infty}|a_{n_j}|^{\varkappa}$$

endlich, $\tilde{p}$ ist offensichtlich eine $\varkappa$-Halbnorm auf $l(F; (p_n))$. Die Zahlenfolge, die an n_j-ter Stelle eine 1, sonst aber nur Nullen hat, bezeichnen wir mit v_j. Die v_j gehören nach Voraussetzung (s. Einleitung) zu F, also auch zu $l(F; (p_n))$. Die Folge (v_j) ist in $l(F; (p_n))$ beschränkt, für jedes n und für beliebige positive Zahlen $\varrho_1, \ldots, \varrho_n$ ist

$$\tilde{p}(\varrho_1 v_1+\cdots+\varrho_n v_n)=\varrho_1^{\varkappa}+\cdots+\varrho_n^{\varkappa}.$$

Nach Satz 2 ist alles bewiesen, d.h. $l(F; (p_n))$ kann nicht s-konvex sein, wenn noch die Stetigkeit von $\tilde{p}$ im Nullelement von $l(F; (p_n))$ gezeigt wird. Hierzu genügt der Nachweis, daß $V=\{x\in l(F; (p_n)): \tilde{p}(x)\leq 1\}$ eine Nullumgebung ist. Es sei nun U die Nullumgebung, die aus allen $x\in l(F; (p_n))$ mit $p(x)\leq 1$ besteht. Für jedes $x=(a_n)$ aus U ist

$$\tilde{p}(x)\leq\sum_{j=1}^{\infty}|a_{n_j}|^{p_{n_j}}\leq p(x)\leq 1.$$

Also ist $V\supset U$, V ist ebenfalls eine Nullumgebung.

Die Räume $L_r(0<r<1)$

Mit kompakt ist bikompakt gemeint. Ist S ein σ-Ring (s. z.B. HALMOS [3]) von Teilmengen einer Menge E mit $\cup\{A: A\in S\}=E$, so soll S ein σ-Ring auf E genannt werden. Ist f eine reellwertige Funktion auf der Menge E, S ein σ-Ring auf E, so heißt f meßbar (HALMOS [3]), wenn für jedes reelle ϱ die Menge $\{x\in E: f(x)\neq 0$ und $f(x)<\varrho\}$ zu S gehört. Damit ist auch klar, was die Meßbarkeit einer komplexwertigen Funktion bedeutet. Bei dieser Definition sind insbesondere diejenigen auf E erklärten charakteristischen Funktionen φ_A meßbar, bei denen $A=\{x\in E: \varphi(x)=1\}$ zu S gehört, d.h. meßbar ist. Ist E ein topologischer Raum, so sei Σ_0 der kleinste σ-Ring auf E, der alle kompakten Teilmengen von E, die ein G_δ sind, enthält. Die Mengen von Σ_0 sind die BAIREschen Mengen von E.

Es sei nun S ein σ-Ring auf einer Menge E, μ ein Maß auf S, Λ_r $(0<r<1)$ der lineare Funktionenraum aller komplexwertigen meßbaren Funktionen f auf E mit integrierbarem $|f|^r$. Alle $f\in\Lambda_r$ mit $\int|f|^r d\mu=0$ bilden einen Teilraum N_r von Λ_r. Die „Quasimetrik"

$$d(f, g)=\int|f-g|^r d\mu \qquad (f, g\in\Lambda_r)$$

definiert in bekannter Weise auf dem Quotientenraum $L_r=\Lambda_r/N_r$ eine Metrik, durch die L_r zu einem linearen topologischen Raum wird.

Satz 6. *E sei ein lokalkompakter Raum mit unendlich vielen Elementen. S sei ein σ-Ring auf E, der den σ-Ring Σ_0 aller* BAIRE*schen Mengen enthält.*

μ sei ein Maß auf S, für jede kompakte BAIREsche *Menge A, die einen inneren Punkt besitzt, sei $\mu(A)$ endlich und positiv. Dann ist L_r $(0<r<1)$ für kein s mit $r<s\leq 1$ s-konvex.*

Beweis. Mit $\emptyset$ wird die leere Menge bezeichnet. Es existiert, wie eine einfache Überlegung zeigt, eine Folge (y_n) von Elementen von E und eine Folge (W_n) von Teilmengen von E, so daß W_n Umgebung von y_n und $W_m \cap W_n = \emptyset$ $(m \neq n)$ ist. Jedes W_n enthält eine Umgebung V_n von y_n, die ein kompaktes G_δ ist (HALMOS [3], S. 218). Weiter existiert zu jedem n eine stetige Abbildung f_n von E in $[0, 1]$ mit $f_n(y_n)=1$, $f_n(x)=0$ $(x \notin V_n)$. f_n ist meßbar (sogar bezüglich Σ_0; HALMOS [3], S. 220), wegen $\mu(V_n)<+\infty$ ist $|f_n|^r$ integrierbar, wegen $\mu(V_n)>0$ ist auch

$$J_n = \int |f_n|^r \, d\mu$$

positiv. Wir setzen $g_n = (1/J_n)^{1/r} f_n$ $(n=1, 2, \ldots)$. Gehört f zu Λ_r, so sei $[f]$ das zugehörige Element in $L_r = \Lambda_r/N_r$. Die $[g_n]$ bilden in L_r eine beschränkte Folge, durch $\psi([f]) = \int |f|^r d\mu$ $(f \in \Lambda_r)$ ist auf L_r eine stetige r-Norm definiert, für jedes n und für beliebige positive Zahlen $\varrho_1, \ldots, \varrho_n$ ist

$$\psi(\varrho_1[g_1] + \cdots + \varrho_n[g_n]) = \varrho_1^r + \cdots + \varrho_n^r.$$

Nach Satz 2 ist damit die Behauptung bewiesen.

Im folgenden sei E eine Gruppe mit dem Einselement e_0, S wiederum ein σ-Ring auf E, μ ein σ-finites Maß auf S. Alle $A \times B$ $(A, B \in S)$ bilden einen σ-Ring $S \times S$ auf $E \times E$, T sei die durch $(x, y) \to (x, xy)$ erklärte Abbildung von $E \times E$ in sich. Ist A eine Teilmenge von E, so sei A' das Komplement von A in E. Für zwei Teilmengen A, B von E wird $A \Delta B = (A \cap B') \cup (A' \cap B)$ gesetzt (HALMOS [3]). E wird als meßbare Gruppe bezeichnet, wenn folgendes gilt: S und μ sind linksinvariant $\big(xA \in S, \mu(xA) = \mu(A)$ für alle $x \in E, A \in S\big)$, die Abbildung T führt die (bezüglich $S \times S$) meßbaren Teilmengen von $E \times E$ wieder in meßbare Mengen über, zu jedem von e_0 verschiedenen $x \in E$ existiert ein $A \in S$ mit $0<\mu(A)<+\infty$, $\mu\big((xA) \Delta A\big) > 0$.

Satz 7. *Die meßbare Gruppe E besitze unendlich viele Elemente. Dann ist L_r $(0<r<1)$ für kein s mit $r<s\leq 1$ s-konvex. Insbesondere ist L_r nicht lokalkonvex.*

Beweis. Für $A \in S$, $0<\mu(A)<+\infty$, $0<\varepsilon<2\mu(A)$ sei

$$U(A; \varepsilon) = \{x \in E : \mu\big((xA) \Delta A\big) < \varepsilon\}.$$

Es gibt dann genau eine Topologie T_0 auf E, bei der E zu einer topologischen Gruppe wird und die $U(A; \varepsilon)$ eine Basis für den Filter aller Umgebungen von e_0 bilden. T_0 ist die sog. WEILsche Topologie von E (HALMOS [3], S. 273). Es existiert dann eine lokalkompakte Gruppe $\hat{E}$, so daß E isomorph zu einer Untergruppe $\tilde{E}$ von $\hat{E}$ ist, die in $\hat{E}$ überall dicht liegt. Für die folgenden Überlegungen kann man $\tilde{E} = E$ annehmen. Wir bezeichnen eine Teilmenge A

von E als präkompakt (vgl. BOURBAKI [1], S. 161), wenn in $\hat{E}$ die abgeschlossene Hülle von A kompakt ist; versehen mit der WEILschen Topologie ist also E eine lokal präkompakte Gruppe.

Die Topologie von $\hat{E}$, die auf E die WEILsche Topologie T_0 erzeugt, nennen wir $\hat{T}_0$. Ist A eine Teilmenge von E, so soll die abgeschlossene Hülle von A in E bzw. in $\hat{E}$ mit $\bar{A}$ bzw. mit A^* bezeichnet werden. Es gibt nun (s. Beweis von Satz 6) eine Folge (y_n) von Elementen aus E und eine Folge (W_n) von Teilmengen von E, so daß W_n T_0-Umgebung von y_n, $W_m \cap W_n = \emptyset$ $(m \neq n)$ und $\overline{W_n} = W_n$ ist. W_n^* ist dann eine $\hat{T}_0$-Umgebung von y_n und enthält eine $\hat{T}_0$-Umgebung V_n von y_n, die (in $\hat{E}$) ein kompaktes G_δ ist. Da $\hat{E}$ vollständig regulär ist, existiert zu jedem n eine stetige Abbildung φ_n von $\hat{E}$ in $[0, 1]$ mit $\varphi_n(y_n) = 1$, $\varphi_n(x) = 0$ $(x \in \hat{E},\ x \notin V_n)$. Alle Teilmengen A von $\hat{E}$ mit $A \cap E \in S$ bilden einen σ-Ring $\hat{S}$ auf $\hat{E}$, der die BAIREschen Mengen von $\hat{E}$ enthält (HALMOS [3], S. 276). Insbesondere gehören also alle V_n zu $\hat{S}$, für jedes n ist daher $V_n \cap E$ meßbar in E. Nach einer oben gemachten Bemerkung (s. Beweis von Satz 6) ist φ_n meßbar (bezüglich $\hat{S}$). f_n sei die Einschränkung von φ_n auf E. Dann ist f_n meßbar bezüglich S und $f_n(x) = 0$ $(x \in E,\ x \notin V_n \cap E)$. Wegen $0 < \mu(V_n \cap E) < +\infty$ (HALMOS [3], S. 274) ist $|f_n|^r$ integrierbar und $\int |f_n|^r d\mu > 0$. Nun ist $V_n \cap E \subset W_n^* \cap E = \overline{W}_n = W_n$, also $(V_m \cap E) \cap (V_n \cap E) = \emptyset$ für $m \neq n$. Der Rest des Beweises verläuft wie bei Satz 6.

Literatur

[1] BOURBAKI, N.: Topologie générale, Chap. I—II. Actual. Sci. Industr. 858—1142 (1951). — [2] DAY, M. M.: The spaces L^p with $0 < p < 1$. Bull. Amer. Math. Soc. **46**, 816—832 (1940). — [3] HALMOS, P. R.: Measure Theory. New York 1950. — [4] LIVINGSTON, A. E.: The space H^p, $0 < p < 1$, is not normable. Pacific J. Math. **3**, 613—616 (1953). — [5] MAZUR, S., et W. ORLICZ: Sur les espaces métriques linéaires. I. Studia Math. **10**, 184—208 (1948). — [6] NAKANO, H.: Concave modulars. J. Math. Soc. Japan **5**, 29—49 (1953). — [7] PÓLYA, G., u. G. SZEGÖ: Aufgaben und Lehrsätze aus der Analysis. I. Berlin 1925. — [8] TYCHONOFF, A.: Ein Fixpunktsatz. Math. Ann. **111**, 767—776 (1935). — [9] WALTERS, S. S.: The space H^p with $0 < p < 1$. Proc. Amer. Math. Soc. **1**, 800—805 (1950). — [10] WALTERS, S. S.: Remarks on the space H^p. Pacific J. Math. **1**, 455—471 (1951).

Dresden, Mathematisches Institut der Technischen Hochschule

(Eingegangen am 3. Januar 1956)

ENDL, K.
Math. Zeitschr. Bd. 65, S. 113—132 (1956)

Über Klassen von Limitierungsverfahren, die die Klasse der Hausdorffschen Verfahren als Spezialfall enthalten

Von
KURT ENDL

Einleitung

Ein HAUSDORFF-Verfahren H ist bekanntlich ein lineares Limitierungsverfahren, das sich mit Hilfe einer beliebigen Diagonalmatrix $\mu = (\mu_n)$ und der Differenzenmatrix $\delta = \left((-)^n \binom{m}{n}\right)$ in der Form

$$H = \delta \mu \delta \tag{1}$$

darstellen läßt. Eine Reihe von wichtigen Eigenschaften der Klasse $\mathfrak{H}$ aller HAUSDORFF-Verfahren beruht auf der Möglichkeit der Darstellung eines jeden Verfahrens in der Form (1) bzw. auf der Tatsache, daß die Differenzenmatrix δ gleich ihrer Inversen ist

$$\delta = \delta^{-1}. \tag{2}$$

So sind z.B. die Verträglichkeit und Vertauschbarkeit aller HAUSDORFF-Verfahren auf (1) und (2) zurückzuführen. Es liegt nun nahe, Klassen von Limitierungsverfahren zu untersuchen, die auf eine zu (1) analoge Weise mit Hilfe von Diagonalmatrizen μ und einer Matrix δ^* mit der Eigenschaft $\delta^* = (\delta^*)^{-1}$ erzeugt werden.

Wir legen unseren Untersuchungen eine abzählbare Klasse von Matrizen $\delta^{(k)}$ $(\delta^{(k)} = (\delta^{(k)})^{-1};\ k = 1, 2, \ldots)$ zugrunde, die in natürlicher Weise die Differenzenmatrix δ verallgemeinern $(\delta = \delta^{(1)})$ und gewinnen auf diese Weise abzählbare viele Klassen $\mathfrak{M}^{(k)}$ von Limitierungsverfahren. Die HAUSDORFFsche Klasse $\mathfrak{H}$ tritt als einfachster Fall auf: $\mathfrak{H} = \mathfrak{M}^{(1)}$. Für jede dieser Klassen $\mathfrak{M}^{(k)}$ existiert ein Verfahren $C^{(k)} \in \mathfrak{M}^{(k)}$ mit der Eigenschaft: Die Klasse $\mathfrak{M}^{(k)}$ besteht aus allen linearen Limitierungsverfahren, die mit $C^{(k)}$ vertauschbar sind. Ferner gelten für alle k die schönen Eigenschaften der HAUSDORFFschen Klasse: Alle Verfahren einer Klasse $\mathfrak{M}^{(k)}$ sind verträglich und vertauschbar. Die Frage nach dem Durchschnitt von verschiedenen Klassen $\mathfrak{M}^{(k)}$ ist vollständig gelöst: Zwei verschiedene Klassen haben nur gewisse triviale Verfahren gemeinsam, deren Matrizen Diagonalmatrizen sind.

Betrachten wir die Unterklassen $\mathfrak{R}^{(k)}$ der regulären Verfahren von $\mathfrak{M}^{(k)}$, so gilt sogar: Die Klassen $\mathfrak{R}^{(k)}$ sind bis auf das Einheitsverfahren paarweise fremd.

Die Verfahren $(M^{(k)}, \mu)$ der Klasse $\mathfrak{M}^{(k)}$ erlauben eine einfache Deutung. Jedes Verfahren $(M^{(k)}, \mu)$ ist identisch mit einem Verfahren, das durch „gleichzeitige Anwendung" von k HAUSDORFF-Verfahren $H_0, H_1, H_2, \ldots, H_{k-1}$

entsteht. Hierbei ist „gleichzeitige Anwendung" in dem Sinne zu verstehen, daß jedes Teilverfahren H_ϱ auf eine Teilfolge $\left\{s_n^{(\varrho)}\right\} = \{s_{\varrho+nk}\}$ der zu limitierenden Folge $\{s_n\}$ anzuwenden ist ($\varrho = 0, 1, \ldots, k-1$). Wir werden in diesem Sinne die neuen Verfahren unmißverständlich als „Produkt" von HAUSDORFF-Verfahren schreiben

$$(M^{(k)}, \mu) = H_0 \times H_1 \times \cdots \times H_{k-1}.$$

Von den gewonnenen Regularitätskriterien sei nur das folgende angeführt: Das Verfahren $(M^{(k)}, \mu) = H_0 \times H_1 \times \cdots \times H_{k-1}$ ist genau dann regulär, wenn jedes Teilverfahren H_ϱ regulär ist ($\varrho = 0, 1, \ldots, k-1$).

Zu Wirkfelduntersuchungen wird in bequemer Weise die geometrische Reihe herangezogen. Während nach AGNEW im Falle der HAUSDORFF-Verfahren H die geometrische Reihe in Kreisen summiert wird, sind für $k \geq 2$ die Limitierungsgebiete der geometrischen Reihe in bezug auf die neuen Verfahren $(M^{(k)}, \mu)$ CASSINIsche Ovale. Dies erlaubt, neben der Bemerkung, daß jetzt im Gegensatz zu den HAUSDORFF-Verfahren die geometrische Reihe für $k \geq 2$ in Punkten der rechten Halbebene $\Re(z) > 1$ summiert wird, folgende interessante Aussage über analytische Fortsetzung: Das BOREL-OKADAsche Gebiet $M^{(k)}(B)$ [dies ist das Gebiet, in dem die Potenzreihen $\sum_0^\infty a_n z^n$ *aller* auf dem Sterngebiet B regulär-analytischen Funktionen $f(z)$ durch $(M^{(k)}, \mu)$ summiert werden] ist für gerade k bei beliebigem B nullpunktsymmetrisch.

Das Problem der Verträglichkeit von Verfahren aus verschiedenen Klassen ist ungelöst. Auch über Vergleichbarkeit bzw. Unvergleichbarkeit von Verfahren verschiedener Klassen ist allgemein noch nichts bekannt. Es sei nur ein Ergebnis hervorgehoben, das vielleicht schon ein wenig Licht auf diese schwierig erscheinenden Fragen wirft. Wir werden aus jeder Klasse $\mathfrak{M}^{(k)}$ ($k = 1, 2, \ldots$) ein Verfahren $M^{*(k)}$ herausgreifen und beweisen: Die Verfahren $M^{*(k)}$ sind für verschiedene k paarweise unvergleichbar.

Schließlich sei noch eine Beobachtung erwähnt, die darauf hinweist, in welcher Weise die neuen Verfahren zur Gewinnung eines Einblicks in die Struktur von Folgen benutzt werden können. Wir geben eine Folge an, die weder $(C, 1)$- noch $(E, 1)$-limitierbar ist, dagegen durch das Verfahren $(C, 1) \times (E, 1)$ zu $\frac{1}{2}$ limitiert wird. Dies bedeutet, daß man durch eine geeignete Zusammenstellung von HAUSDORFF-Verfahren $H_0, H_1, \ldots, H_{k-1}$ zu einem unserer Produktverfahren $H_0 \times H_1 \times \cdots \times H_{k-1}$ Folgen limitieren kann, für welche die einzelne Anwendung jedes Verfahrens versagt.

§1. HAUSDORFF-Verfahren[1])

Wir stellen hier kurz die Eigenschaften der HAUSDORFF-Verfahren zusammen, auf die wir im Verlauf der Arbeit fortwährend zurückgreifen müssen.

[1]) Alle hier angeführten Sätze über HAUSDORFF-Verfahren findet man z. B. in dem Lehrbuch von G. H. HARDY: Divergent Series (Oxford University Press) in Kapitel XI, Hausdorff Means (247—282). Man findet dort auch ein Literaturverzeichnis.

Definition der HAUSDORFF-*Verfahren*

Ein HAUSDORFF-Verfahren ist ein lineares Limitierungsverfahren

$$t_m = \sum_{n=0}^{m} h_{mn} s_n ,$$

dessen Matrix $H = (h_{mn})$ in der Form $H = \delta \mu \delta$ darstellbar ist, wobei δ die Differenzenmatrix $\left((-)^n \binom{m}{n}\right)$ und μ eine beliebige Diagonalmatrix (μ_n) ist. Wir sprechen im folgenden kurz von einem H-Verfahren und bezeichnen es mit (H, μ) oder, wenn keine Verwechslung zu befürchten ist, mit H. Dieselben Symbole bedeuten gleichzeitig auch die zugehörigen Matrizen.

Die CAESAROschen Verfahren (C, k), die EULERschen Verfahren (E, q) und die HÖLDERschen Verfahren (H, k) sind H-Verfahren.

Es gelten die folgenden Sätze:

Satz 1. Zwei H-Verfahren H_1 und H_2 sind vertauschbar.

Satz 2. Alle H-Verfahren sind verträglich[2]).

Satz 3. Die Klasse $\mathfrak{H}$ der H-Verfahren ist identisch mit der Klasse aller linearen Verfahren, die mit dem $(C, 1)$-Verfahren vertauschbar sind.

Ein Limitierungsverfahren A heißt regulär oder permanent, wenn jede Folge $\{s_n\}$, die im gewöhnlichen Sinn zu s konvergiert, durch das Verfahren A ebenfalls zum Wert s limitiert wird. Bezeichnen wir eine Folge $\{a_n\}$ als total monoton, wenn für alle $n = 0, 1, \ldots$ und $p = 0, 1, \ldots$ gilt

$$\Delta^p a_n = \sum_{\nu=0}^{p} (-)^\nu \binom{p}{\nu} a_{n+\nu} \geq 0 ,$$

so kann das Regularitätskriterium für H-Verfahren folgendermaßen formuliert werden:

Satz 4. Das Verfahren (H, μ) ist genau dann regulär, wenn die Folge $\{\mu_n\}$ Differenz zweier total monotoner Folgen ist, $\Delta^m \mu_0 \to 0$ strebt und $\mu_0 = 1$ ist.

Ist $\chi(x)$ auf $[0, 1]$ eine Funktion von beschränkter Schwankung, so heißt

$$\mu_n = \int_0^1 x^n \, d\chi$$

die Momentenkonstante der Funktion χ vom Rang n, $\{\mu_n\}$ die zu χ gehörige Momentenfolge. Gilt noch weiter $\chi(1) = 1$ und $\chi(+0) = \chi(0) = 0$ (stetiges Verschwinden am 0-Punkt), so heißt die Momentenkonstante μ_n bzw. die Momentenfolge $\{\mu_n\}$ regulär. Definieren wir schließlich den Funktionswert von χ an Unstetigkeitsstellen $x \in (0, 1)$ durch $\chi(x) = \frac{1}{2}[\chi(x-0) + \chi(x+0)]$, so gilt: Die Funktion $\chi(x)$ ist durch die reguläre Folge $\{\mu_n\}$ eindeutig bestimmt. Ein reguläres Verfahren (H, μ) ist hiernach auch durch die zu der regulären Momentenfolge $\{\mu_n\}$ gehörigen Funktion $\chi(x)$ bestimmt. Wir schreiben

[2]) Zwei Verfahren A und B heißen verträglich, wenn für jede Folge $\{s_n\}$ die von A zu dem Wert s, von B zum Wert s' limitiert wird, gilt $s = s'$.

deshalb für reguläre Verfahren statt (H, μ) manchmal auch (H, χ). Der Zusammenhang zwischen Momentenfolgen und total monotonen Folgen wird durch die folgenden Sätze gegeben:

Satz 5. Jede Momentenfolge ist Differenz zweier total monotoner Folgen und umgekehrt.

Satz 6. Für jede reguläre Momentenfolge $\{\mu_n\}$ gilt $\mu_0 = 1$ und $\Delta\mu_0 \to 0$.

Aus Satz 4 und Satz 6 folgt damit eine andere Form des Regularitätskriteriums:

Satz 7. Das Verfahren (H, μ) ist genau dann regulär, wenn $\{\mu_n\}$ eine reguläre Momentenfolge ist.

Das allgemeine Matrixelement h_{mn} des Verfahrens (H, μ) drückt sich in den Folgenelementen μ_n folgendermaßen aus

$$h_{mn} = \binom{m}{n} \Delta^{m-n} \mu_n. \tag{1.1}$$

Ist (H, μ) regulär, so gilt mit der zugehörigen Funktion $\chi(x)$ auch die Darstellung

$$h_{mn} = \binom{m}{n} \int_0^1 (1-x)^{m-n} x^n \, d\chi. \tag{1.2}$$

Über den Vergleich von regulären H-Verfahren $H = (H, \mu)$ und $H' = (H, \mu')$, für deren Diagonalmatrizen $\mu = (\mu_n)$ und $\mu' = (\mu'_n)$ für alle n $\mu_n \neq 0$ und $\mu'_n \neq 0$ ist, gilt

Satz 8. Das reguläre Verfahren (H, μ) ist genau dann stärker als das reguläre Verfahren (H, μ'), wenn die Folge $\left\{\frac{\mu_n}{\mu'_n}\right\}$ eine reguläre Momentenfolge ist.

Korollar. Die Verfahren (H, μ) und (H, μ') sind genau dann äquivalent, wenn sowohl $\left\{\frac{\mu_n}{\mu'_n}\right\}$ als auch $\left\{\frac{\mu'_n}{\mu_n}\right\}$ eine reguläre Momentenfolge ist.

§2. Einführung der verallgemeinerten Differenzenmatrizen $\delta^{(k)}$

Wir führen jetzt eine abzählbare Klasse von Matrizen ein, die in natürlicher Weise die Differenzenmatrix $\delta = \left((-)^n \binom{m}{n}\right)$ verallgemeinert.

Sei k eine feste natürliche Zahl. Jede andere natürliche Zahl m läßt sich in der Form darstellen $m = m_k^* + m_k \cdot k$ wobei $0 \leq m_k^* < k$ ist. Wir definieren unter Benutzung des Symbols

$$\delta_{mn}^{(k)} = \begin{cases} 1 & \text{für } m \equiv n\,(k) \\ 0 & \text{für } m \not\equiv n\,(k) \end{cases}$$

die Matrix

$$\delta^{(k)} = \left((-)^{n_k} \binom{m_k}{n_k} \delta_{m^*_k n^*_k}^{(k)}\right)$$

und behaupten:

Satz 8'. Für jede natürliche Zahl k ist $\delta^{(k)} = (\delta^{(k)})^{-1}$ oder $(\delta^{(k)})^2 = \mathfrak{E}$ ($\mathfrak{E}$ Einheitsmatrix).

Beweis. Das Element in der m-ten Zeile und n-ten Spalte der Produktmatrix $\delta^{(k)} \cdot \delta^{(k)}$

$$(2.1) \qquad \sum_{l=n}^{m} (-)^{l_k} \binom{m_k}{l_k} \delta_{ml}^{(k)} \cdot (-)^{n_k} \binom{l_k}{n_k} \cdot \delta_{ln}^{(k)} = (-)^{n_k} \binom{m_k}{n_k} \sum_{l=n}^{m} (-)^{l_k} \binom{m_k - n_k}{l_k - n_k} \delta_{ml}^{(k)} \delta_{ln}^{(k)}$$

verschwindet immer für $m \not\equiv n\,(k)$, da nur über die l summiert wird, für die gleichzeitig $l \equiv m\,(k)$ und $l \equiv n\,(k)$ gilt. Ist $m \equiv n\,(k)$, so liefern genau die Werte $l = n, n+k, \ldots n+(m_k - n_k)k$ einen von 0 verschiedenen Beitrag zur Summe (2.1). Wir können dafür also schreiben

$$(-)^{n_k} \binom{m_k}{n_k} \sum_{\lambda=n_k}^{m_k} (-)^{\lambda} \binom{m_k - n_k}{\lambda - n_k} = \binom{m_k}{n_k} \sum_{\lambda=0}^{m_k - n_k} (-)^{\lambda} \binom{m_k - n_k}{\lambda} = \begin{cases} 1 & n_k = m_k, \\ 0 & n_k \neq m_k \end{cases}$$

w. z. b. w.

Für $k = 1$ ist $\delta_{mn}^{(1)} = 1$ für beliebige m und n. Da ferner immer $m_1 = m$ ist, erhalten wir $\delta^{(1)} = \left((-)^n \binom{m}{n}\right) = \delta$, die gewöhnliche Differenzenmatrix.

§3. Definition und Deutung der neuen Verfahren

Wir benutzen die in § 2 eingeführten Matrizen $\delta^{(k)}$ $(k = 1, 2, 3, \ldots)$ zur Erzeugung von linearen Limitierungsverfahren, indem wir wie in der Hausdorffschen Theorie mit einer beliebigen Diagonalmatrix $\mu = (\mu_n)$ das Produkt $(M^{(k)}, \mu) = \delta^{(k)} \mu\, \delta^{(k)}$ bilden. Wieder soll auch für das durch die Matrix $(M^{(k)}, \mu)$ gegebene Verfahren dieselbe Bezeichnung $(M^{(k)}, \mu)$ verwendet werden. Zu jedem k erhalten wir eine Klasse von Limitierungsverfahren $\mathfrak{M}^{(k)}$, indem wir für μ alle Diagonalmatrizen zulassen. $\mathfrak{M}^{(1)}$ ist die Klasse der H-Verfahren. Bevor wir auf innere Eigenschaften der Klassen $\mathfrak{M}^{(k)}$ und den Vergleich verschiedener Klassen eingehen, sollen die Matrixelemente $\mu_{mn}^{(k)}$ der Matrix $(M^{(k)}, \mu)$ berechnet werden.

Aus der Definition $(M^{(k)}, \mu) = \delta^{(k)} \mu\, \delta^{(k)}$ folgt:

$$\mu_{mn}^{(k)} = \sum_{l=n}^{m} (-)^{l_k} \binom{m_k}{l_k} \delta_{ml}^{(k)} \cdot \mu_l \cdot (-)^{n_k} \binom{l_k}{n_k} \delta_{ln}^{(k)}$$

und daraus wegen der Bedeutung der Symbole $\delta_{ml}^{(k)}$, $\delta_{ln}^{(k)}$

$$(3.1) \qquad \mu_{mn}^{(k)} = 0 \qquad \text{für } m \not\equiv n\,(k).$$

Ist $m \equiv n\,(k)$, so erhalten wir durch analoge Überlegungen wie in § 1

$$(3.2) \qquad \mu_{mn}^{(k)} = \binom{m_k}{n_k} \sum_{\lambda=0}^{m_k - n_k} (-)^{\lambda} \binom{m_k - n_k}{\lambda} \mu_{n+\lambda k}.$$

Setzen wir zur Abkürzung $\Delta_k^p \mu_q = \sum_{\lambda=0}^{p} (-)^{\lambda} \binom{p}{\lambda} \mu_{q+\lambda k}$ $(p, q \geq 0)$, so kann (3.1) und (3.2) in der folgenden Form zusammengefaßt werden.

$$(3.3) \qquad \mu_{mn}^{(k)} = \binom{m_k}{n_k} \delta_{mn}^{(k)} \cdot \Delta_k^{m_k - n_k} \mu_n.$$

Da $\delta_{m,n}^{(1)} = 1$ ist für alle m und n und stets $m_1 = m$ gilt, ist (3.3) für $k = 1$ identisch mit (1.1), der Darstellung des HAUSDORFF-Elementes.

Die Struktur der Verfahren $(M^{(k)}, \mu)$ erlaubt eine einfache Deutung. Für die Transformierte

$$t_m = \sum_{n=0}^{m} \mu_{m,n}^{(k)} s_n$$

folgt nämlich aus (3.3)

$$t_m = \sum_{\nu=0}^{m_k} \mu_{m,m_k+\nu k}^{(k)} s_{m_k+\nu k}. \tag{3.4}$$

Betrachten wir alle m in derselben Restklasse $\varrho\ (0 \leq \varrho < k)$, d.h. alle m von der Form $\varrho + \lambda k$ $(\lambda = 0, 1, \ldots)$, so lautet (3.4)

$$t_{\varrho+\lambda k} = \sum_{\nu=0}^{\lambda} \mu_{\varrho+\lambda k,\varrho+\nu k}^{(k)} \cdot s_{\varrho+\nu k}. \tag{3.5}$$

Dies bedeutet aber, daß die Teilfolge $\{s_{\varrho+\nu k}\} = \{s_\nu^{(\varrho)}\}$ der Folge $\{s_n\}$ in die Teilfolge $\{t_{\varrho+\nu k}\} = \{t_\nu^{(\varrho)}\}$ der Folge $\{t_n\}$ transformiert wird. Das Verfahren $(M^{(k)}, \mu)$ enthält also k Teilverfahren T_ϱ, von denen jedes genau die Teilfolge $\{s_\nu^{(\varrho)}\}$ in die Teilfolge $\{t_\nu^{(\varrho)}\}$ transformiert. Wir behaupten: Die Teiltransformation T_ϱ ist ein HAUSDORFF-Verfahren; seine Diagonalmatrix lautet $\mu^{(\varrho)} = \{\mu_n^{(\varrho)}\} = \{\mu_{\varrho+nk}\}$.

Beweis. Aus (3.5) und (3.3) entnehmen wir

$$t_\lambda^{(\varrho)} = \sum_{\nu=0}^{\lambda} \mu_{\varrho+\lambda k,\varrho+\nu k}^{(k)} s_\nu^{(\varrho)} = \sum_{\nu=0}^{\lambda} \binom{\lambda}{\nu} \Delta_k^{\lambda-\nu} \mu_{\varrho+\nu k} \cdot s_\nu^{(\varrho)}.$$

Da

$$\Delta_k^{\lambda-\nu} \mu_{\varrho+\nu k} = \sum_{\sigma=0}^{\lambda-\nu} (-)^\sigma \binom{\lambda-\nu}{\sigma} \mu_{\varrho+\nu k+\sigma k}$$

$$= \sum_{\sigma=0}^{\lambda-\nu} (-)^\sigma \binom{\lambda-\nu}{\sigma} \mu_{\nu+\sigma}^{(\varrho)} = \Delta^{\lambda-\nu} \mu_\nu^{(\varrho)} \quad (\mu_n^{(\varrho)} = \mu_{\varrho+nk})$$

ist, folgt

$$t_\lambda^{(\varrho)} = \sum_{\nu=0}^{\lambda} \binom{\lambda}{\nu} \Delta^{\lambda-\nu} \mu_\nu^{(\varrho)} \cdot s_\nu^{(\varrho)}.$$

Dies ist aber gerade die zu der Matrix $\mu^{(\varrho)} = (\mu_n^{(\varrho)})$ gehörige H-Transformation. Wir formulieren diesen Sachverhalt in dem

Satz 9. Jedes Verfahren $M^{(k)} = (M^{(k)}, \mu)$ ist identisch mit k „gleichzeitig" angewandten HAUSDORFF-Verfahren $H_\varrho = (H, \mu^{(\varrho)})$ $(\mu^{(\varrho)} = (\mu_n^{(\varrho)}) = (\mu_{\varrho+nk}))$. Hierbei ist „gleichzeitig" in dem Sinn zu verstehen, daß jedes Teilverfahren H_ϱ genau auf die Teilfolge $\{s_n^{(\varrho)}\} = \{s_{\varrho+nk}\}$ anzuwenden ist $(\varrho = 0, 1, \ldots, k-1)$. Wir werden das Verfahren $M^{(k)}$ unmißverständlich als „Produkt" der k Teilverfahren H_ϱ schreiben

$$M^{(k)} = H_0 \times H_1 \times \cdots \times H_{k-1}.$$

§4. Regularitätskriterien

Wir untersuchen, für welche Diagonalmatrizen $\mu=(\mu_n)$ die Verfahren $(M^{(k)},\mu)$ regulär sind.

Nach TOEPLITZ sind die notwendigen und hinreichenden Bedingungen für die Regularität eines Verfahrens $t_m=\sum_{n=0}^{\infty} c_{mn}\,s_n$:

(I) $\gamma_m=\sum_{n=0}^{\infty}|c_{mn}|<K$ unabhängig von m.

(II) $c_{mn}\to 0$ für $m\to\infty$ und n beliebig.

(III) $c_m=\sum_{n=0}^{\infty}c_{mn}\to 1$ für $m\to\infty$.

Beginnen wir mit (III). Da für eine beliebige Folge $\{\lambda_m\}$ stets

$$\sum_{n=0}^{m}\binom{m}{n}\Delta^{m-n}\lambda_n=\lambda_0$$

ist[3]), folgt für unsere Verfahren, wenn wir alle m in derselben Restklasse $\varrho\,(0,1,\ldots,k-1)$ betrachten

$$c_m=\sum_{n=0}^{m}\mu_{mn}^{(k)}=\sum_{\lambda=0}^{m_k}\binom{m_k}{\lambda}\Delta_k^{m_k-\lambda}\,\mu_{\varrho+\lambda k}=\sum_{\lambda=0}^{m_k}\binom{m_k}{\lambda}\Delta^{m_k-\lambda}\,\mu_\lambda^{(\varrho)}=\mu_0^{(\varrho)}.$$

Setzen wir also $\mu_0^{(\varrho)}=1$ für $\varrho=0,1,\ldots,k-1$, d.h. $\mu_0=\mu_1=\cdots=\mu_{k-1}=1$, so gilt offenbar für beliebiges m $c_m=1$, d.h. (III) ist erfüllt. Ist auch nur ein $\mu_0^{(\varrho)}\neq 1$, so gilt für die ganze Teilfolge $c_{\varrho+\lambda k}\neq 1$, d.h. c_m konvergiert nicht gegen 1. Die Bedingung $\mu_0=\mu_1=\cdots=\mu_{k-1}=1$ ist also äquivalent mit (III).

Wir nehmen nun an, daß (I) für das Verfahren $(M^{(k)},\mu)$ erfüllt ist, d.h. gilt

$$\gamma_m=\sum_{\lambda=0}^{m_k}\binom{m_k}{\lambda}\left|\Delta_k^{m_k-\lambda}\,\mu_{m_k^2+\lambda k}\right|<K \tag{4.1}$$

unabhängig von m. Dann gilt dies insbesondere auch für alle m in derselben Restklasse ϱ

$$\gamma_{\varrho+\lambda k}=\sum_{\sigma=0}^{\lambda}\binom{\lambda}{\sigma}\left|\Delta_k^{\lambda-\sigma}\,\mu_{\varrho+\sigma k}\right|=\sum_{\sigma=0}^{\lambda}\binom{\lambda}{\sigma}\left|\Delta^{\lambda-\sigma}\,\mu_\sigma^{(\varrho)}\right|<K \tag{4.2}$$

für alle λ und $\varrho=0,1,\ldots,k-1$. Dies ist aber nach HAUSDORFF[4]) äquivalent mit der Aussage: Die Folge $\{\mu_\lambda^{(\varrho)}\}$ ist Differenz zweier total monotoner Folgen

$$\{\mu_\lambda^{(\varrho)}\}=\{\alpha_\lambda^{(\varrho)}-\beta_\lambda^{(\varrho)}\},\quad \Delta^\varkappa\alpha_\lambda^{(\varrho)}\geq 0,\quad \Delta^\varkappa\beta_\lambda^{(\varrho)}\geq 0\qquad(\lambda,\varkappa=0,1,\ldots).$$

Da ϱ eine beliebige der Zahlen $0,1,\ldots,k-1$ war, gilt, wenn wir $\alpha_\lambda^{(\varrho)}=\alpha_{\varrho+\lambda k}$ und $\beta_\lambda^{(\varrho)}=\beta_{\varrho+\lambda k}$ setzen für die Folgen $\{\alpha_n\}$ und $\{\beta_n\}$:

$$\Delta_k^m\alpha_n\geq 0\quad\text{und}\quad\Delta_k^m\beta_n\geq 0\qquad m,n=0,1,\ldots.$$

---

[3]) Vgl. 1., S. 252.

[4]) Vgl. 1., S. 253—254.

Nennen wir solche Folgen total monoton von der Ordnung k (total monoton von der Ordnung 1 $\sim$ total monoton), so folgt also aus (4.1): Die Folge $\{\mu_n\}$ ist Differenz zweier von der Ordnung k total monotoner Folgen $\{\mu_n\} = \{\alpha_n - \beta_n\}$. Ist umgekehrt $\{\mu_n\}$ Differenz zweier von der Ordnung k total monotoner Folgen $\{\alpha_n\}$ und $\{\beta_n\}$, d.h. jede der Folgen $\{\alpha_\lambda^{(\varrho)}\}$ und $\{\beta_\lambda^{(\varrho)}\}$ $(\varrho = 0, 1, \ldots, k-1)$ total monoton, so gilt nach dem eben erwähnten Satz von HAUSDORFF für die Folge $\{\mu_\lambda^{(\varrho)}\} = \{\alpha_\lambda^{(\varrho)} - \beta_\lambda^{(\varrho)}\}$ die Ungleichung (4.2) mit einer Konstanten K_ϱ. Wählen wir $K = \max_{\varrho=0}^{k-1} \{K_\varrho\}$, so folgt (4.1) für alle m mit derselben Konstanten K. Die TOEPLITZsche Bedingung (I) ist also für unsere Verfahren äquivalent mit der Bedingung: Die Folge $\{\mu_n\}$ ist Differenz zweier von der Ordnung k total monotoner Folgen.

Die Bedingung (II) lautet hier

$$c_{m\,n} = \mu_{m\,n}^{(k)} = \binom{m_k}{n_k} \delta_{m\,n}^{(k)} \Delta_k^{m_k - n_k} \mu_n \to 0 \quad \text{für } m \to \infty \text{ und } n \text{ beliebig,}$$

oder für eine beliebige Restklasse ϱ

$$(4.3)\quad c_{\varrho+\lambda k,\,\varrho+\nu k} = \mu_{\varrho+\lambda k,\,\varrho+\nu k}^{(k)} = \binom{\lambda}{\nu} \Delta^{\lambda-\nu} \mu_\nu^{(\varrho)} \to 0 \quad \text{für } \lambda \to \infty \text{ und } \nu \text{ beliebig.}$$

Nun folgt aber (4.3) nach einem bekannten Satz von HAUSDORFF[5]) schon aus

$$\sum_{\sigma=0}^{\lambda} \binom{\lambda}{\sigma} |\, \Delta^{\lambda-\sigma} \mu_\sigma^{(\varrho)} \,| < K$$

für alle $\nu > 0$, so daß also unter Voraussetzung der Gültigkeit von (4.1) die Bedingung (II) äquivalent mit den k Bedingungen ist

$$\Delta^\lambda \mu_0^{(\varrho)} = \Delta_k^\lambda \mu_\varrho \to 0 \qquad \text{für } \lambda \to \infty \text{ und } \varrho = 0, 1, \ldots, k-1.$$

Wir können diese Ergebnisse zu dem Kriterium zusammenfassen

Satz 10. Das Verfahren $M^{(k)} = (M^{(k)}, \mu)$ ist genau dann regulär, wenn

(I') die Folge $\{\mu_n\}$ Differenz zweier von der Ordnung k total monotoner Folgen ist,

(II') $\mu_\varrho = 1$ ist und

(III') $\Delta_k^m \mu_\varrho \to 0$ für $m \to \infty$ $\Big\}$ $\varrho = 0, 1, 2, \ldots, k-1.$

Da die Aussage: $\{\mu_n\}$ ist Differenz zweier von der Ordnung k total monotoner Folgen gleichbedeutend damit ist, daß jede der k Teilfolgen $\{\mu_n^{(\varrho)}\}$ Differenz zweier total monotoner Folgen ist, und da ferner $\{\mu_n^{(\varrho)}\}$ gerade die zu dem ϱ-ten Teilverfahren $H_\varrho = (H, \mu^{(\varrho)})$ gehörige Folge ist (Satz 7), ist das obige Kriterium nach Satz 9 äquivalent mit der folgenden Fassung

Satz 11. Das Verfahren $M^{(k)} = H_0 \times H_1 \times H_2 \times \cdots \times H_{k-1}$ ist genau dann regulär, wenn jedes der k Teilverfahren H_ϱ regulär ist $(\varrho = 0, 1, \ldots, k-1)$.

[5]) Vgl. 1., S. 255.

§5. Bemerkung zum Momentenproblem

Ist das Verfahren $M^{(k)} = (M^{(k)}, \mu) = H_0 \times \cdots \times H_{k-1}$ regulär, so auch jedes Teilverfahren $H_\varrho = (H, \mu^{(\varrho)})$ (Satz 11). Dies ist aber nach Satz 7 äquivalent mit der Aussage: Die Folge $\{\mu_n^{(\varrho)}\} = \{\mu_{\varrho+nk}\}$ ist eine reguläre Momentenfolge. Es existiert also (unter der Voraussetzung $M^{(k)}$ regulär) auf $[0, 1]$ für jedes ϱ eine Funktion χ_ϱ von beschränkter Schwankung und normalen Unstetigkeiten, die den Bedingungen $\chi_\varrho(1) = 1$, $\chi_\varrho(+0) = \chi_\varrho(0) = 0$ genügt und die Darstellung erlaubt

$$\mu_n^{(\varrho)} = \int_0^1 x^n \, d\,\chi_\varrho \qquad \varrho = 0, 1, \ldots, k-1 .$$

Unter Benutzung des schon mehrfach benutzten Symbols $\delta_{mn}^{(k)}$ können wir mit diesen Funktionen χ_ϱ die Folge $\{\mu_n\}$ in der Form darstellen

$$\mu_n = \sum_{\varrho=0}^{k-1} \delta_{n\,\varrho}^{(k)} \int_0^1 x^{n_k} \, d\,\chi_\varrho .$$

Nennen wir eine Folge $\{\mu_n\}$ für welche ein solcher Satz von k Funktionen χ_ϱ gefunden werden kann, eine reguläre Momentenfolge der Ordnung k, so können wir unser Regularitätskriterium auch in der Form aussprechen.

Satz 12. Das Verfahren $M^{(k)} = (M^{(k)}, \mu)$ ist genau dann regulär, wenn $\{\mu_n\}$ eine reguläre Momentenfolge der Ordnung k ist.

§6. Bestimmung und Eigenschaften der Klassen $\mathfrak{M}^{(k)}$

Wir betrachten für ein beliebiges fest gewähltes k die Klasse $\mathfrak{M}^{(k)}$ aller Verfahren $(M^{(k)}, \mu) = \delta^{(k)} \mu \, \delta^{(k)}$ und beweisen als erstes

Satz 13. Je zwei Verfahren der Klasse $\mathfrak{M}^{(k)}$ sind vertauschbar.

Beweis. Sind $\mu = (\mu_n)$ und $\mu' = (\mu'_n)$ zwei beliebige Diagonalmatrizen, so gilt wegen der Vertauschbarkeit zweier Diagonalmatrizen und der Beziehung $(\delta^{(k)})^2 = \mathfrak{E}$

$$\begin{aligned}(M^{(k)}, \mu)\,(M^{(k)}, \mu') &= \delta^{(k)} \mu \, \delta^{(k)} \cdot \delta^{(k)} \mu' \, \delta^{(k)} = \delta^{(k)} \mu \, \mu' \, \delta^{(k)} \\ &= \delta^{(k)} \mu' \, \mu \, \delta^{(k)} = \delta^{(k)} \mu' \, \delta^{(k)} \cdot \delta^{(k)} \, \mu \, \delta^{(k)} = (M^{(k)}, \mu')\,(M^{(k)}, \mu) .\end{aligned}$$

Insbesondere sind alle Verfahren $(M^{(k)}, \mu)$ mit dem Verfahren $C^{(k)} = (M^{(k)}, \mu^*)$ vertauschbar, wo $\mu_n^* = \dfrac{1}{n+1}$ ist. Wir bezeichnen die zugehörigen Teiltransformationen T_ϱ mit $C_\varrho^{(k)}$ und setzen also $C^{(k)} = C_0^{(k)} \times C_1^{(k)} \times \cdots \times C_{k-1}^{(k)}$. Als nächstes wollen wir nun beweisen, daß die Klasse $\mathfrak{M}^{(k)}$ in dem Sinn „abgeschlossen" ist, daß nicht nur alle Verfahren dieser Klasse mit $C^{(k)}$ vertauschbar sind, sondern auch umgekehrt jedes lineare Verfahren $A: t_m = \sum_{n=0}^{m} a_{mn} s_n$, das mit $C^{(k)}$ vertauschbar ist, schon Element der Klasse $\mathfrak{M}^{(k)}$ ist. Dazu zeigen wir zuerst, daß aus der Voraussetzung der Vertauschbarkeit mit $C^{(k)} = (c_{mn}^{(k)})$

$$(a_{mn})\,(c_{mn}^{(k)}) = (c_{mn}^{(k)})\,(a_{mn}) \tag{6.1}$$

die folgende Struktureigenschaft über die Matrix (a_{mn}) folgt:

$$a_{mn} = 0 \qquad \text{für } m \not\equiv n(k). \tag{6.2}$$

Beweis. Wir nehmen zuerst an, daß die natürlichen Zahlen m und n $(m \not\equiv n(k))$ den Bedingungen genügen $m-k<n<m$. Unter dieser Voraussetzung ergibt sich für (6.1) (wir setzen für $n<0$ $a_{mn}=0$)

$$\sum_{\substack{i=n \\ i\equiv n(k)}}^{m} a_{mi}\, c_{in}^{(k)} = \sum_{\substack{i=n \\ i\equiv m(k)}}^{m} c_{mi}^{(k)}\, a_{in} \tag{6.3}$$

und damit $a_{mn}\, c_{nn}^{(k)} = c_{mm}^{(k)}\, a_{mn}$, d.h. aber $a_{mn}=0$, da ja nach Wahl des Verfahrens $C^{(k)}$ für $n \not\equiv m$ $c_{nn}^{(k)} \not\equiv c_{mm}^{(k)}$ ist.

Wir machen nun die Induktionsannahme, daß für alle natürlichen Zahlen m und n $(m \not\equiv n(k))$ mit $m-\nu\cdot k<n<m$ das Element $a_{mn}=0$ ist (ν natürliche Zahl).

Behauptung. Dies gilt dann auch für alle natürlichen Zahlen m und n $(m \not\equiv n(k))$ mit $m-(\nu+1)k<n<m$.

Aus (6.1) entnehmen wir

$$a_{mn}\, c_{nn}^{(k)} + \sum_{\substack{i=n+k \\ i\equiv n(k)}}^{m} a_{mi}\, c_{in}^{(k)} = c_{mm}^{(k)}\, a_{mn} + \sum_{\substack{i=n \\ i\equiv m(k)}}^{m-k} c_{mi}^{(k)} \cdot a_{in}.$$

Da auf beiden Seiten unter den Summenzeichen nur a_{mn} auftreten, die nach Induktionsvoraussetzung 0 sind, folgt $a_{mn}\cdot c_{nn}^{(k)} = c_{mm}^{(k)}\, a_{mn}$ und hieraus wieder wegen der Wahl von $C^{(k)}$ $a_{mn}=0$. Damit ist (6.2) allgemein bewiesen.

Das Verfahren A besitzt also dieselbe Struktur wie jedes Verfahren der Klasse $\mathfrak{M}^{(k)}$.

$$t_m = \sum_{\lambda=0}^{m_k} a_{m,\, m_k^* + \lambda k} \cdot s_{m_k^* + \lambda k}.$$

Hieraus schließen wir, daß sich A wieder als Produkt von k Teilverfahren A_ϱ schreiben läßt, wobei A_ϱ die Folge $\{s_n^{(\varrho)}\} = \{s_{\varrho+nk}\}$ in die Teilfolge $\{t_n^{(\varrho)}\} = \{t_{\varrho+nk}\}$ transformiert;

$$A = A_0 \times A_1 \times \cdots \times A_{k-1}.$$

Dann ist aber, wie man sofort aus der Struktur der beiden Matrizen $C^{(k)}$ und A folgert, die Vertauschbarkeit von $C^{(k)}$ und A äquivalent mit der Vertauschbarkeit jedes der Teilverfahren $C_\varrho^{(k)}$ mit A_ϱ $(\varrho = 0, 1, \ldots, k-1)$. Nach einem Satz von HAUSDORFF[6]) ist dann schon A_ϱ ein H-Verfahren, d.h. es existiert eine Diagonalmatrix $\mu^{(\varrho)} = (\mu_n^{(\varrho)})$, so daß A_ϱ in der Form $A_\varrho = \delta\mu^{(\varrho)}\delta$ geschrieben werden kann. Hieraus folgt aber, wenn wir $\mu_n^{(\varrho)} = \mu_{\varrho+nk}$ und $\mu = (\mu_n)$ setzen, für das Gesamtverfahren A die Darstellung $A = \delta^{(k)}\, \mu\, \delta^{(k)}$, d.h. $A \in \mathfrak{M}^{(k)}$. Es gilt also

[6]) Vgl. 1., S. 249.

Satz 14. Die Klasse $\mathfrak{M}^{(k)}$ besteht aus allen linearen Verfahren, die mit $C^{(k)}$ vertauschbar sind.

Als eine weitere wichtige Eigenschaft der Klasse $\mathfrak{M}^{(k)}$ beweisen wir

Satz 15. Alle Verfahren einer Klasse $\mathfrak{M}^{(k)}$ sind verträglich.

Beweis. Die Aussage, daß die beiden Verfahren $(M^{(k)}, \mu) = H_0 \times H_1 \times \cdots \times H_{k-1}$ und $(M^{(k)}, \mu') = H_0' \times H_1' \times \cdots \times H_{k-1}'$ die Folge $\{s_n\}$ zu dem Grenzwert s bzw. s' limitieren, bedeutet: Jede der k Teilfolgen $\{s_n^{(\varrho)}\}$ wird von H_ϱ zu s bzw. von H_ϱ' zu s' limitiert. Da je zwei H-Verfahren verträglich sind, folgt $s = s'$.

§7. Vergleich von verschiedenen Klassen $\mathfrak{M}^{(k)}$

Eine interessante Frage[6']) ist naturgemäß die nach dem Durchschnitt von verschiedenen Klassen $\mathfrak{M}^{(k)}$. Sind zwei Klassen $\mathfrak{M}^{(k)}$ und $\mathfrak{M}^{(k')}$ für $k \neq k'$ vielleicht sogar fremd? Wir zeigen, daß die Klassen $\mathfrak{M}^{(k)}$ und $\mathfrak{M}^{(k')}$ nur gewisse triviale Verfahren gemeinsam haben, die durch Diagonalmatrizen gegeben werden.

Die Aussage $A \in \mathfrak{M}^{(k)}$ und $A \in \mathfrak{M}^{(k')}$ bedeutet die folgenden Möglichkeiten der Darstellung von A

$$A = \delta^{(k)}\, \mu\, \delta^{(k)} \quad \text{und} \quad A = \delta^{(k')}\, \mu'\, \delta^{(k')}, \tag{7.1}$$

wobei μ und μ' eventuell verschiedene Diagonalmatrizen sind. Da aber nach (3.3) für die Diagonalglieder $a_{n\,n} = \mu_n = \mu_n'$ folgt, gilt $\mu = \mu'$. Allgemein zieht (7.1) nach (3.3) die Gleichung nach sich

$$\binom{m_k}{n_k} \Delta_k^{m_k - n_k} \mu_n\, \delta_{m\,n}^{(k)} = \binom{m_{k'}}{n_{k'}} \Delta_{k'}^{m_{k'} - n_{k'}} \mu_n \cdot \delta_{m\,n}^{(k')}. \tag{7.2}$$

Ist nun z.B. $k < k'$, so wählen wir für alle $m \geq k$ $n = m - k$. Für diese m und n ist $\delta_{m\,n}^{(k')} = 0$, d.h. wir haben

$$\binom{m_k}{m_k - 1} \Delta_k^1\, \mu_{m-k} = 0$$

und daraus

$$m_k \sum_{\lambda=0}^{1} (-)^\lambda \binom{1}{\lambda} \mu_{m-k+\lambda k} = m_k (\mu_{m-k} - \mu_m) = 0,$$

d.h. $\mu_m = \mu_{m-k}$. Dies bedeutet aber, daß jede der k Folgen $\{\mu_\lambda^{(\varrho)}\} = \{\mu_{\varrho+\lambda k}\}$ $(\lambda = 0, 1, \ldots;\ \varrho = 0, 1, \ldots, k-1)$ konstant ist. Dann gilt aber schon für alle m und n

$$a_{m\,n} = \binom{m_k}{n_k} \delta_{m\,n}^{(k)} \Delta_k^{m_k - n_k} \mu_n = \binom{m_k}{n_k} \delta_{m\,n}^{(k)} \sum_{\lambda=0}^{m_k - n_k} (-)^\lambda \binom{m_k - n_k}{\lambda} \mu_{n+\lambda k}$$

$$= \binom{m_k}{n_k} \delta_{m\,n}^{(k)} \cdot \mu_\varrho \cdot \delta_{m_k\, n_k} = \begin{cases} \mu_\varrho & m = n \\ 0 & m \neq n \end{cases}$$

[6']) Ein Problem dieser Art ist von E. Ullrich für Haussdorff- und Nörlund-Verfahren behandelt worden. Math. Z. 25 (1926), 382—387.

d.h. A ist eine Diagonalmatrix. Da jetzt die linke Seite von (7.2) immer 0 ist, wenn $m \neq n$, folgt genau so als zweite Strukturbedingung der Folge $\{\mu_n\}$: $\mu_{\varrho+\lambda k'} = \mu_\varrho$ $(\lambda = 0, 1, \ldots;\ \varrho = 0, 1, \ldots, k'-1)$. Liegt umgekehrt eine Diagonalmatrix $A = (\mu_n)$ vor, für welche $\mu_\varrho = \mu_{\varrho+\lambda k}$ $(\lambda = 0, 1, \ldots;\ \varrho = 0, 1, \ldots, k-1)$ bzw. $\mu_\varrho = \mu_{\varrho+\lambda k'}$ $(\lambda = 0, 1, 2, \ldots;\ \varrho = 0, 1, \ldots, k'-1)$ ist, so gilt sowohl $A = \delta^{(k)} \mu\, \delta^{(k)}$ als auch $A = \delta^{(k')} \mu\, \delta^{(k')}$ mit $\mu = (\mu_n)$, d.h. es gilt $A \in \mathfrak{M}^{(k)}$ und $A \in \mathfrak{M}^{(k')}$. Damit folgt

Satz 16. Die Klassen $\mathfrak{M}^{(k)}$ und $\mathfrak{M}^{(k')}$ sind für $k \neq k'$ fremd bis auf gewisse triviale Verfahren, deren Matrizen Diagonalmatrizen $\mu = (\mu_n)$ sind. Die Folgen $\{\mu_n\}$ müssen dabei noch zusätzlich den Strukturbedingungen $\mu_\varrho = \mu_{\varrho+\lambda k}$ $(\lambda = 0, 1, \ldots;\ \varrho = 0, 1, \ldots, k-1)$ und $\mu_\varrho = \mu_{\varrho+\lambda k'}$ $(\lambda = 0, 1, \ldots;\ \varrho = 0, 1, \ldots, k'-1)$ genügen.

§8. Die Unterklassen $\mathfrak{R}^{(k)}$ der regulären Verfahren von $\mathfrak{M}^{(k)}$

Wir betrachten jetzt die Unterklassen $\mathfrak{R}^{(k)}$ der regulären Verfahren von $\mathfrak{M}^{(k)}$. Über den Durchschnitt zweier Unterklassen $\mathfrak{R}^{(k)}$ und $\mathfrak{R}^{(k')}$ können wir sogar beweisen

Satz 17. Die Unterklassen $\mathfrak{R}^{(k)}$ und $\mathfrak{R}^{(k')}$ $(k \neq k')$ haben genau ein Verfahren gemeinsam, nämlich das Einheitsverfahren.

Beweis. Nach Satz 16 ist ein gemeinsames Verfahren notwendig ein Diagonalverfahren (μ_n), dessen Diagonalfolge $\{\mu_n\}$ noch zusätzlich den Bedingungen genügen muß: $\mu_\varrho = \mu_{\varrho+\lambda k}$ $(\lambda = 0, 1, \ldots;\ \varrho = 0, 1, \ldots, k-1)$. Da nach dem Regularitätskriterium Satz 10 für reguläre Verfahren

$$\mu_0 = \mu_1 = \cdots = \mu_{k-1} = 1$$

ist, folgt hieraus $\mu_n = 1$ $(n = 1, 2, \ldots)$ w. z. b. w.

Über Vergleichbarkeit von Verfahren einer Klasse beweisen wir

Satz 18. Genau dann ist $(M^{(k)}, \mu) = H_0 \times H_1 \times \cdots \times H_{k-1} \in \mathfrak{R}^{(k)}$ stärker als $(M^{(k)}, \mu') = H_0' \times H_1' \times \cdots \times H_{k-1} \in \mathfrak{R}^{(k)}$, wenn jedes der Teilverfahren H_ϱ stärker als das entsprechende Teilverfahren H_ϱ' ist $(\varrho = 0, 1, \ldots, k-1)$.

Beweis. Die Beziehung $(M^{(k)}, \mu) \geqq (M^{(k)}, \mu')$ bedeutet, daß jede Folge $\{s_n\}$, die von $(M^{(k)}, \mu')$ limitiert wird, auch von $(M^{(k)}, \mu)$ limitiert wird. Dies bedeutet aber nichts anderes, als daß jede Teilfolge $\{s_n^{(\varrho)}\}$, die von H_ϱ' limitiert wird, auch von H_ϱ limitiert wird, d.h. daß $H_\varrho \geqq H_\varrho'$ ist $(\varrho = 0, 1, \ldots, k-1)$. Die Umkehrung ist evident.

Korollar. Die Verfahren $(M^{(k)}, \mu) \in \mathfrak{R}^{(k)}$ und $(M^{(k)}, \mu') \in \mathfrak{R}^{(k)}$ sind genau dann äquivalent, wenn jedes Teilverfahren H_ϱ mit H_ϱ' äquivalent ist $(\varrho = 0, 1, \ldots, k-1)$.

Gilt für die Matrizen $\mu = (\mu_n)$ und $\mu' = (\mu_n')$ $\mu_n \neq 0$, $\mu_n' \neq 0$ $(n = 0, 1, 2, \ldots)$, so erkennen wir die Gültigkeit von

Satz 19. Das Verfahren $(M^{(k)}, \mu) \in \mathfrak{R}^{(k)}$ ist genau dann stärker, als das Verfahren $(M^{(k)}, \mu') \in \mathfrak{R}^{(k)}$, wenn die Folge $\left\{\frac{\mu_n}{\mu_n'}\right\}$ eine reguläre Momentenfolge der Ordnung k ist.

Beweis. $(M^{(k)}, \mu) \geqq (M^{(k)}, \mu')$ bedeutet $H_\varrho \geqq H'_\varrho$ $(\varrho = 0, 1, \ldots, k-1)$. Nach Satz 7 ist dann $\left\{\frac{\mu_n^{(\varrho)}}{\mu'^{(\varrho)}_n}\right\}$ eine reguläre Momentenfolge, woraus nach Definition einer regulären Momentenfolge der Ordnung k die Behauptung folgt.

§9. Summation der geometrischen Reihe in CASSINIschen Ovalen und analytische Fortsetzung

Sehr oft können aus dem Summationsgebiet der geometrischen Reihe in bezug auf ein Limitierungsverfahren A Aussagen über die Güte, d.h. die Größe des Wirkfeldes von A, gemacht werden. Wir werden sehen, daß wir auch hier die geometrische Reihe in sehr bequemer Weise zur Entscheidung von Vergleichbarkeit bzw. Nichtvergleichbarkeit von Verfahren aus verschiedenen Klassen $\mathfrak{M}^{(k)}$ heranziehen können. Ferner wird sich für unsere Verfahren aus einer besonderen Eigenschaft des Summationsgebietes der geometrischen Reihe eine interessante Folgerung für die EULERschen Summierungspolygone ergeben.

Wir benutzen zur Berechnung des Summationsgebietes der geometrischen Reihe in bezug auf ein Verfahren $(M^{(k)}, \mu) = H_0 \times H_1 \times \cdots \times H_{k-1}$ einen Satz von AGNEW über H-Verfahren[7]). Dieser Satz lautet:

Ist $H(\chi)$ ein beliebiges reguläres H-Verfahren und verstehen wir unter r die untere Grenze aller r', für welche gilt $\chi(t) = 1$ wenn $r' \leq t \leq 1$, so wird die geometrische Reihe $\sum_0^\infty z^n$ durch $H(\chi)$ in jedem inneren Punkt des Kreises $z - \left(1 - \frac{1}{r}\right) = \frac{1}{r}$ zu $\frac{1}{1-z}$ summiert; dagegen ist $\sum_{\nu=0}^\infty z^\nu$ in keinem äußeren Punkt dieses Kreises durch $H(\chi)$ summierbar.

Alle H-Verfahren summieren also die geometrische Reihe in Kreisen, die durch $z = 1$ gehen und deren Mittelpunkt auf der negativen reellen Achse liegt. Wir nennen mit AGNEW r die Ordnung des H-Verfahrens (H, χ) und den Kreis $z - \left(1 - \frac{1}{r}\right) = \frac{1}{r}$ den Summierungskreis von (H, χ).

Wir skizzieren für eine spätere Anwendung den Beweis von AGNEW. Die (H, χ)-Transformierte der Folge $\{s_n\}$ lautet nach (1.2) in Integraldarstellung

$$t_m = \sum_{n=0}^{m} h_{m\,n}\, s_n = \int_0^1 \left\{ \sum_{n=0}^{m} \binom{m}{n} x^n (1-x)^{m-n}\, s_n \right\} d\chi .$$

Damit erhält man für die Teilsummen der geometrischen Reihe $s_n(z) = 1 + z + \cdots + z^n$

$$t_m(z) = \frac{1}{1-z} - \frac{z}{1-z} \int_0^1 [1 + t(z-1)]^n \, d\chi .$$

[7]) AGNEW, R. P.: Analytic Extension by Hausdorff Methods. Trans. Amer. Math. Soc. **52**, 217—237 (1942).

AGNEW zeigt nun, daß die Funktion $f_n(z) = \int_0^1 [1+t(z-1)]^n d\chi$ [die (H, χ)-Transformierte der Folge $\{z^n\}$] für jeden Punkt innerhalb des Kreises $\left|z-\left(1-\frac{1}{r}\right)\right| = \frac{1}{r}$ gegen 0 konvergiert, für jeden Punkt außerhalb unbeschränkt ist.

Nach § 5 gilt für die Folge $\{\mu_n\}$ eines beliebigen regulären Verfahrens $(M^{(k)}, \mu)$ die Momentendarstellung

$$\mu_n = \sum_{\varrho=0}^{k-1} \delta_{n\varrho}^{(k)} \int_0^1 x^{n_k} d\chi_\varrho,$$

wo χ_ϱ zu der regulären Momentenfolge $\{\mu_{\varrho+nk}\} = \{\mu_n^{(\varrho)}\}$ gehört $(\varrho = 0, 1, \ldots, k-1)$. Das allgemeine Matrixelement schreibt sich dann in der Form

$$\begin{aligned}
\mu_{mn}^{(k)} &= \binom{m_k}{n_k} \Delta_k^{m_k-n_k} \mu_n \cdot \delta_{mn}^{(k)} \\
&= \binom{m_k}{n_k} \delta_{mn}^{(k)} \cdot \sum_{\lambda=0}^{m_k-n_k} (-)^\lambda \binom{m_k-n_k}{\lambda} \mu_{n+\lambda k} \\
&= \binom{m_k}{n_k} \delta_{mn}^{(k)} \sum_{\lambda=0}^{m_k-n_k} (-)^\lambda \binom{m_k-n_k}{\lambda} \cdot \sum_{\varrho=0}^{k-1} \delta_{n+\lambda k,\varrho}^{(k)} \int_0^1 x^{n_k+\lambda} d\chi_\varrho \\
&= \binom{m_k}{n_k} \delta_{mn}^{(k)} \int_0^1 x^{n_k} (1-x)^{m_k-n_k} d\chi_{n_k^*}.
\end{aligned}$$

Angewandt auf die Teilsumme der geometrischen Reihe $s_n(z) = 1 + z + \cdots + z^n = \frac{1-z^{n+1}}{1-z}$ erhalten wir:

$$\begin{aligned}
t_m(z) &= \sum_{n=0}^{m} \mu_{mn}^{(k)} s_n(z) \\
&= \frac{1}{1-z} \int_0^1 (1-x)^{m_k} \left\{ \sum_{\lambda=0}^{m_k} \binom{m_k}{\lambda} \frac{x^\lambda}{(1-x)^\lambda} \right\} d\chi_{m_k^*} \\
&\qquad - \frac{z^{m_k^*+1}}{1-z} \cdot \int_0^1 (1-x)^{m_k} \left\{ \sum_{\lambda=0}^{m_k} \binom{m_k}{\lambda} \left(\frac{x}{1-x}\right)^\lambda (z^k)^\lambda \right\} d\chi_{m_k^*} \\
&= \frac{1}{1-z} - \frac{z^{m_k^*+1}}{1-z} \int_0^1 \left(1 - x(1-z^k)\right)^{m_k} d\chi_{m_k^*}.
\end{aligned}$$

Wir betrachten nun für sich die k Teilfolgen $\{t_m^{(\varrho)}(z)\} = \{t_{\varrho+mk}(z)\}$ $(\varrho = 0, 1, \ldots, k-1)$, für welche gilt

$$t_m^{(\varrho)}(z) = \frac{1}{1-z} - \frac{z^{\varrho+1}}{1-z} \int_0^1 [1 - x(1-z^k)]^m d\chi_\varrho.$$

Setzen wir für den Augenblick $z^k = \zeta$, so ist

$$R_m^{(\varrho)}(z) = \int_0^1 [1 - x(1 - z^k)]^m \, d\chi_\varrho = \int_0^1 [1 - x(1 - \zeta)]^m \, d\chi_\varrho = f_m^{(\varrho)}(\zeta)$$

die (H, χ_ϱ)-Transformierte der Folge $\{\zeta^n\}$. Diese konvergiert nach AGNEW, wenn r_ϱ die Ordnung von (H, χ_ϱ) bezeichnet, in jedem Punkt innerhalb des Kreises $\left|\zeta - \left(1 - \frac{1}{r_\varrho}\right)\right| = \frac{1}{r_\varrho}$ gegen 0 und ist unbeschränkt in jedem Punkt außerhalb. Bezeichnen wir $r = \operatorname{Max}_{\varrho=0}^{k-1} \{r_\varrho\}$ als die Ordnung des Verfahrens $(M^{(k)}, \mu)$ so konvergiert im Inneren des Gebietes $\left|\zeta - \left(1 - \frac{1}{r}\right)\right| = \frac{1}{r}$ jede der k Teilfolgen $\{t_m^{(\varrho)}(z)\}$ zu $\frac{1}{1-z}$; in jedem Punkt außerhalb bleibt dagegen wenigstens eine dieser Folgen unbeschränkt. Damit haben wir bewiesen:

Satz 20. Besitzt das Verfahren $(M^{(k)}, \mu) = H_0 \times H_1 \times \cdots \times H_{k-1}$ $(k \geq 2$; $H_\varrho = (H, \mu_n^{(\varrho)}) = (H, \chi_\varrho))$ die Ordnung r, so summiert es die geometrische Reihe in jedem Punkt innerhalb des CASSINIschen Ovals $\left|z^k - \left(1 - \frac{1}{r}\right)\right| = \frac{1}{r}$ zu $\frac{1}{1-z}$ dagegen in keinem Punkt außerhalb.

Als besonders erwähnenswert heben wir noch die Tatsache hervor, daß es in den Klassen $\mathfrak{M}^{(k)}$ $(k \geq 2)$ Verfahren gibt, die die geometrische Reihe in Punkten der rechten Halbebene $\Re(z) > 1$ summieren. Dies ist für kein HAUSDORFF-Verfahren der Fall.

Zur Behandlung der analytischen Fortsetzung von Potenzreihen stützen wir uns auf einen Satz von BOREL-LORENTZ-OKADA[8]). Wir betrachten ein beliebiges Verfahren $A: t_m = \sum a_{mn} s_n$, das die geometrische Reihe innerhalb[9]) eines offenen Sterngebietes[10]) R', das den Einheitskreis enthält, gleichmäßig zu $\frac{1}{1-z}$ summiert. Mit R bezeichnen wir die aus $C(R')$ (Komplement von R') durch die Stürzung $w = \frac{1}{z}$ erhaltene Menge. Diese ist nach unseren obigen Voraussetzungen über R' ein abgeschlossenes Sterngebiet und enthält das Intervall $[0, 1]$. Ist nun B eine beliebige Menge, die 0 enthält, $c \cdot C$ für eine beliebige Menge C die Menge aller Punkte cz mit $z \in C$ $(c \neq 0)$ und bezeichnen wir mit $A(B)$ das Gebiet, in dem die Potenzreihe $\sum_0^\infty a_\nu z^\nu$ einer jeden, in B regulär analytischen Funktion $f(z)$ zu $f(z)$ summiert wird, so gilt

1. $A(B) = \bigcap_{c \notin B} cR' = \bigcup_{cR \subseteq B} \{c\}$.

2. $A(B)$ ist sternförmig.

3. Die Potenzreihe einer jeden in B analytischen Funktion $f(z)$ ist innerhalb B gleichmäßig zu $f(z)$ summierbar.

[8]) LORENTZ, G. G.: Bernstein Polynomials, S. 117—120. University of Toronto Press 1953.

[9]) Innerhalb eines Gebietes bedeutet auf jeder abgeschlossenen Teilmenge.

[10]) Ein Sterngebiet enthält den 0-Punkt und auf jedem Strahl von 0 aus genau einen Randpunkt.

Dieser Satz darf offenbar für unsere Verfahren $(M^{(k)}, \mu)$ $(k \geq 2)$ angewandt werden. Denn R' ist in diesem Fall ein CASSINIsches Oval, das den Einheitskreis enthält und sternförmig ist. Ferner strebt nach AGNEW die Funktion $f_n^{(\varrho)}(\zeta)$ innerhalb des zu (H, χ_ϱ) gehörigen Limitierungskreises gleichmäßig gegen 0 und damit die Funktionenfolge $t_m(z)$ innerhalb des zu dem Durchschnitt dieser Limitierungskreise $(\varrho = 0, 1, \ldots, k-1)$ gehörigen CASSINIschen Ovals gleichmäßig gegen $\frac{1}{1-z}$.

Eine interessante Beobachtung ergibt sich aus der Bemerkung, daß das CASSINIsche Oval $z^k - \left(1 - \frac{1}{r}\right) = \frac{1}{r}$ für gerade k nullpunktsymmetrisch ist. Da die Stürzung $w = \frac{1}{z}$ ein nullpunktsymmetrisches Gebiet wieder in ein nullpunktsymmetrisches Gebiet überführt, erhält man aus der Definition von $A(B)$

Satz 21. Ist k gerade und B eine beliebige vorgegebene Menge die den Nullpunkt enthält, so summiert jedes Verfahren $(M^{(k)}, \mu)$ die Potenzreihe $\sum_{\nu=0}^{n} a_\nu z^\nu$ aller in B regulär analytischen Funktionen $f(z)$ in einem sternförmigen, nullpunktsymmetrischen Gebiet.

§10. Vergleich von Verfahren aus verschiedenen Klassen $\mathfrak{M}^{(k)}$

Besonders interessante Fragen ergeben sich bei der Betrachtung und dem Vergleich von Verfahren aus verschiedenen Klassen $\mathfrak{M}^{(k)}$. Sind z. B. zwei Verfahren aus verschiedenen Klassen immer verträglich oder nie verträglich? Gibt es darunter vergleichbare Verfahren oder nicht? Diese Fragen sind noch offen. Wir zeigen nur an einem Beispiel die Existenz von abzählbar vielen Verfahren aus verschiedenen Klassen, die paarweise unvergleichbar sind.

Wir legen ein beliebiges reguläres HAUSDORFF-Verfahren $(H, \chi^*) = H^*$ mit einer Ordnung $r^* < 1$ zugrunde (d. h. H^* summiert die geometrische Reihe in einem Kreise, der den Einheitskreis als echte Teilmenge enthält) und bilden die Verfahren

$$M^{*(k)} = \underbrace{H^* \times H^* \times \cdots \times H^*}_{k\text{-mal}}.$$

Natürlich ist $M^{*(k)} \in \mathfrak{R}^{(k)}$. Für diese abzählbar vielen Verfahren beweisen wir:

Satz 22. Von den Verfahren $M^{*(k)}$ $(k = 1, 2, \ldots)$ ist keines stärker oder schwächer als ein anderes. Sie sind paarweise unvergleichbar.

Der Beweis ergibt sich unter Heranziehung unserer Ergebnisse über die Summierung der geometrischen Reihe. Nach Satz 20 summiert $M^{*(k)}$ die geometrische Reihe in jedem Punkt des durch die Kurve $z^k - \left(1 - \frac{1}{r^*}\right) = \frac{1}{r^*}$ begrenzten Gebietes $\mathfrak{L}_k$. Wir zeigen nun, daß es zu jedem k Punkte von $\mathfrak{L}_k$ gibt, die in keinem anderen $\mathfrak{L}_{k'}$ liegen $(k' \neq k)$.

Das CASSINIsche Oval $z^k-\left(1-\frac{1}{r^*}\right)=\frac{1}{r^*}$ $(k\geq 2)$ entsteht durch die konforme Abbildung $z=\sqrt[k]{\zeta}$ aus dem Kreis $\zeta-\left(1-\frac{1}{r^*}\right)=\frac{1}{r^*}$. Dieser Kreis mit dem Mittelpunkt $1-\frac{1}{r^*}$ auf der negativen reellen Achse genügt in Polarkoordinaten (r,φ) der Gleichung

$$r=r^*(\varphi)=-\left(\frac{1}{r^*}-1\right)\cos\varphi+\sqrt{\left(\frac{1}{r^*}-1\right)^2\cos^2\varphi+\frac{2}{r^*}-1}.$$

Das CASSINIsche Oval $z^k-\left(1-\frac{1}{r^*}\right)=\frac{1}{r^*}$ besitzt demnach in Polarkoordinaten die Darstellung

$$r=r^k(\varphi)=\sqrt[k]{r^*(k\varphi)}=\sqrt[k]{-\left(\frac{1}{r^*}-1\right)\cos k\varphi+\sqrt{\left(\frac{1}{r^*}-1\right)^2\cos^2\varphi+\frac{2}{r^*}-1}}.$$

Wir betrachten $r^{(k)}(\varphi)$ zuerst für kleine φ und benutzen die TAYLOR-Entwicklungen für $\sin\varphi$, $\cos\varphi$ und $(1+x)^\alpha$. Aus $\cos\varphi=1-\frac{\varphi^2}{2!}+o(\varphi^3)$ und $\sin\varphi=\varphi+o(\varphi^2)$, $\sin^2\varphi=\varphi^2+o(\varphi^3)$ folgt mit $\frac{1}{r^*}=R$

$$(10.1)\quad\begin{cases} r^*(\varphi)=-(R-1)\left(1-\frac{\varphi^2}{2}+o(\varphi^3)\right)+R\sqrt{1-\left(\frac{R-1}{R}\right)^2(\varphi^2+o(\varphi^3)} \\ \quad=-(R-1)\left(1-\frac{\varphi^2}{2}+o(\varphi^3)\right)+R\left[1-\frac{1}{2}\left(\frac{R-1}{R}\right)^2\{\varphi^2+o(\varphi^3)\}+\right. \\ \quad\left.+o\left(\varphi^2+o(\varphi^3)\right)\right] \\ \quad=1+\frac{R-1}{2R}\varphi^2+o(\varphi^2) \end{cases}$$

und hieraus durch nochmalige Anwendung der binomischen Reihe

$$(10.2)\qquad r^{(k)}(\varphi)=1+\frac{R-1}{2R}\cdot k\cdot\varphi^2+o(\varphi^2).$$

Ist nun $k'>k$, so ist für φ genügend klein $r^{(k')}(\varphi)>r^{(k)}(\varphi)$. Denn setzen wir in diese Ungleichung die Entwicklung (10.2) ein, so folgt

$$1+\frac{R-1}{2R}k'\cdot\varphi^2+o(\varphi^2)>1+\frac{R-1}{2R}\cdot k\varphi^2+o(\varphi^2)$$

und hieraus $k'+o(1)>k+o(1)$, eine Ungleichung, die für genügend kleine φ sicher erfüllt ist. Damit haben wir bewiesen, daß es für $k'>k$ Punkte gibt, die nicht in $\mathfrak{L}_k$ wohl aber in $\mathfrak{L}_{k'}$ liegen. Wir zeigen aber auch das Umgekehrte: Für $k'>k$ gibt es Punkte, die in $\mathfrak{L}_k$ aber nicht in $\mathfrak{L}_{k'}$ liegen, mit anderen Worten: Es gibt Werte $\bar{\varphi}$, für welche

$$(10.3)\qquad r^{(k)}(\varphi)>r^{(k')}(\varphi)\quad\text{ist.}$$

Behauptung. Ein solcher Wert ist $\bar{\varphi}=\frac{\pi}{k}$. Wir verifizieren für $\bar{\varphi}=\frac{\pi}{k}$ die behauptete Ungleichung (10.3) sofort für alle k' $(>k$ beliebig) von der

Form $k'=(2n+1)\cdot k$ $(n=1, 2, \ldots)$. Denn aus der Behauptung

$$\sqrt[k]{-(R-1)\cos\left(k\frac{\pi}{k}\right)+\sqrt{(R-1)^2\cos^2\left(k\cdot\frac{\pi}{k}\right)+2R-1}} > \sqrt[k']{(R-1)\cos\left(k'\frac{\pi}{k}\right)+\sqrt{(R-1)^2\cos^2\left(k'\frac{\pi}{k}\right)+2R-1}}$$

folgt für diese k' die richtige Ungleichung

$$\sqrt[k]{2R-1} > \sqrt[k']{2R-1}\,. \tag{10.4}$$

Ist k' $(>k)$ verschieden von den Zahlen $k(2n+1)$ $(n=1, 2, \ldots)$, so beweisen wir sogar die Ungleichung

$$\sqrt[k']{2R-1} > \sqrt[k']{(R-1)\cos\left(\frac{k'}{k}\cdot\pi\right)+\sqrt{(R-1)^2\cos^2\left(\frac{k'}{k}\cdot\pi\right)+2R-1}}\,. \tag{10.5}$$

die wegen (10.4) sofort (10.3) nach sich zieht. Durch Erheben der beiden Seiten von (10.5) in die k-te Potenz erhält man nacheinander

$$2R-1 > (R-1)\cos\left(\frac{k'}{k}\cdot\pi\right)+\sqrt{(R-1)^2\cos^2\left(\frac{k'}{k}\cdot\pi\right)+2R-1}$$

$$(2R-1)^2-(2R-1) > 2(2R-1)(R-1)\cdot\cos\frac{k'}{k}\cdot\pi$$

und wegen $R-1>0 : 1>\cos\frac{k'}{k}\pi$. Die letzte Ungleichung ist sicher immer dann erfüllt, wenn $\frac{k'}{k}\neq 2n+1$ ist, was aber vorausgesetzt war. Damit ist unsere Behauptung über $\bar{\varphi}=\frac{\pi}{k}$ für alle natürlichen k und k' mit $k'>k$ bewiesen.

§11. Anpassung von Limitierungsverfahren an die Struktur der zu untersuchenden Folgen

Wir wollen die Wirksamkeit unserer neuen Verfahren noch an einem anderen Beispiel[11]) beleuchten. Aus den vielen Fragen, die sich bei der Untersuchung der Wirkfelder von Verfahren verschiedener Klassen ergeben, sei die folgende herausgegriffen: Gibt es zu zwei vorgegebenen HAUSDORFF-Verfahren H und H' eine Folge, die weder von H noch H' summiert wird, dagegen durch $H\times H'$ oder $H'\times H$. Wir wählen $H=(C, 1)$, $H'=(E, 1)$ und zeigen für die Folge

$$s_n^* = \begin{cases} \sqrt{\frac{n}{2}} & \text{für } n=2m^2 \quad (m=0, 1, \ldots) \\ \frac{3}{2}\cdot\frac{1}{3}\left(1-(-2)^{\frac{n-1}{2}+1}\right) & \text{für } n=2m+1 \\ 0 & \text{sonst.} \end{cases}$$

[11]) Dieses Beispiel verdanke ich Herrn A. PEYERIMHOFF, dem ich an dieser Stelle herzlich für sein liebenswürdiges Interesse an dieser Arbeit danken möchte.

Die Folge $\{s_n^*\}$ ist weder $(C, 1)$- noch $(E, 1)$-limitierbar; dagegen wird sie durch das Verfahren $(C, 1)\times(E, 1)$ zum Wert $\frac{1}{2}$ limitiert.

Beweis. Die Teilfolge $s_0^*\, s_2^*\, s_4^* \ldots$ ist die von Herrn KNOPP[12]) angegebene Folge $0\,1\,0\,0\,2\,0\,0\,0\,0\,3\,0 \ldots$, die durch das $(C, 1)$-Verfahren zu $\frac{1}{2}$ limitiert wird, dagegen nicht $(E, 1)$-limitierbar ist.

Die Teilfolge $s_1^*\, s_3^*\, s_5^* \ldots$ ist die mit $\frac{3}{2}$ multiplizierte Folge der Teilsummen der geometrischen Reihe $\sum_0^\infty z^n$ an der Stelle $z=-2$, die also von $(E, 1)$ zu $\frac{3}{2}\cdot\frac{1}{3}=\frac{1}{2}$ limitiert wird, von $(C, 1)$ dagegen nicht erfaßt wird. Hieraus folgt schon aus der Definition unserer neuen Verfahren: Das Verfahren $(C, 1)\times(E, 1)$ limitiert die Folge $\{s_n^*\}$ zu $s^*=\frac{1}{2}$.

$(C, 1)$ auf $\{s_n^*\}$ angewandt liefert die Folge

$$t_n^* = \frac{s_0^* + s_1^* + \cdots + s_n^*}{n+1} = \frac{s_0^* + s_2^* + \cdots + s_{2\left[\frac{n}{2}\right]}^*}{n+1} + \frac{s_1^* + s_3^* + \cdots + s_{2\left[\frac{n+1}{2}\right]-1}^*}{n+1}$$

$$= \frac{\left[\frac{n}{2}\right]+1}{n+1}\cdot\frac{s_0^* + s_2^* + \cdots + s_{2\left[\frac{n}{2}\right]}^*}{\left[\frac{n}{2}\right]+1} + \frac{s_1^* + s_3^* + \cdots + s_{2\left[\frac{n+1}{2}\right]-1}^*}{\left[\frac{n+1}{2}\right]}\cdot\frac{\left[\frac{n+1}{2}\right]}{n+1}.$$

Da die Folgen $\frac{\left[\frac{n}{2}\right]+1}{n+1}$ und $\frac{\left[\frac{n+1}{2}\right]}{n+1}$ beide gegen $\frac{1}{2}$ konvergieren,

$$\frac{s_1^* + s_3^* + \cdots + s_{2\left[\frac{n+1}{2}\right]-1}^*}{\left[\frac{n+1}{2}\right]}$$

aber nach KNOPP nicht konvergiert, ist $\{s_n^*\}$ nicht $(C, 1)$-limitierbar. Ähnlich schließt man durch eine leichte Abschätzung, daß $\{s_n^*\}$ nicht $(E, 1)$-limitierbar ist.

An diesem Beispiel sehen wir, daß es möglich ist, durch eine geeignete Zusammenstellung von H-Verfahren $H_0, H_1, \ldots, H_{k-1}$ zu einem unserer Produktverfahren $H_0\times H_1\times\cdots\times H_{k-1}$ Folgen zu limitieren, für welche die Anwendung jedes einzelnen Verfahrens versagt. Die neuen Verfahren erlauben also, durch geeignete Wahl der Faktoren H_ϱ das Limitierungsverfahren der Struktur der einzelnen Folgen anzupassen.

Die systematische Untersuchung der hier eingeführten Verfahrensklassen $\mathfrak{M}^{(k)}$ und ihrer Beziehungen untereinander ist sicher nicht ohne Interesse. Daneben darf aber vielleicht noch auf etwas anderes aufmerksam gemacht

[12]) KNOPP, K.: Theorie und Anwendung der unendlichen Reihen, S. 532. Berlin: Springer 1931.

werden. Die hier untersuchten Verfahren legen es nahe, allgemein Verfahren $A=(a_{mn})$ von der Struktur

$$t_m=\sum_{\nu=0}^{\infty} a_{m,m_k^*+\nu k}\cdot s_{m_k^*+\nu k}$$

zu betrachten (k natürlich, $m=m_k^*+\nu k$, $0\leq m_k^*<k$), d. h. Verfahren, die man vielleicht als Lückenverfahren der Ordnung k bezeichnen kann. Wie wir gesehen haben, arbeiten solche Verfahren in der Weise, daß k Teilverfahren A_ϱ auf die k Teilfolgen $\{s_{\varrho+\nu k}\}$ $(\varrho=0,1,\ldots,k-1;\ \nu=0,1,\ldots)$ für sich angewandt werden. Wir schrieben das Verfahren als „Produkt" der k Teilverfahren $A=A_0\times A_1\times\cdots\times A_{k-1}$. Es ist evident, daß man so unendlich viele Klassen $\mathfrak{A}^{(k)}$ von Verfahren erhält $(\mathfrak{A}^{(k)}\supset\mathfrak{M}^{(k)})$. Die Matrizen dieser Lückenverfahren erhält man durch eine Abbildung des Raumes aller unendlichen Matrizen $A=(a_{mn})$ in sich: Benutzen wir das Symbol

$$\delta_{mn}^{(k)}=\begin{cases}1 & m\equiv n\,(k)\\ 0 & m\not\equiv n\,(k)\,,\end{cases}$$

so können wir für jedes k diese Abbildung in der Form schreiben:

$$A=(a_{mn})\to A^{(k)}=(\delta_{mn}^{(k)}a_{mn})\,.$$

Das Symbol $\delta_{mn}^{(k)}$, in dieser Form angewandt, erzeugt also durch eine Projektion im Matrizenraum die Klasse $\mathfrak{A}^{(k)}$.

Gießen, Mathematisches Institut der Universität

(Eingegangen am 3. Oktober 1955)

SCHUBERT, H.
Math. Zeitschr. Bd. 65, S. 133—170 (1956)

Knoten mit zwei Brücken

Von
HORST SCHUBERT

Brückendarstellungen von Knoten und Verkettungen wurden von TORRES und FOX [10] und dem Verfasser [6] unabhängig voneinander eingeführt. Da der Kreis der einzige Knoten ist, der eine solche Darstellung mit nur einer Brücke besitzt, liegt es nahe, diejenigen Knoten und Verkettungen zu klassifizieren, die Brückendarstellungen mit zwei Brücken besitzen. Dies sind gerade diejenigen, die früher als *Viergeflechte* Interesse gefunden haben[1]).

Es seien E_1 und E_2 zwei parallele Ebenen im 3-dimensionalen euklidischen Raum R^3. Wir bezeichnen sie als waagerecht und E_1 als die obere von beiden. Eine *Brückendarstellung mit zwei Brücken* besteht aus je zwei Streckenzügen in E_1 und E_2 und vier dazu senkrechten Strecken, die sich zu einer oder zwei orientierten Knotenlinien zusammensetzen.

Ist k eine solche Brückendarstellung, so läßt sich durch Deformation von k eine *Normalform* herstellen, die bei senkrechter Projektion nur endlich viele Doppelpunkte besitzt und bei der die beiden Streckenzüge auf E_2 von den Projektionen der auf E_1 liegenden Streckenzüge abwechselnd getroffen werden, sofern k nicht zwei unverkettete Kreise darstellt (§ 3). Sieht man von den Endpunkten der Streckenzüge ab, so erhält man Unterkreuzungen, deren Anzahl von der Form $2n$ ist. Falls $n \neq 0$ ist, also Unterkreuzungen vorhanden sind, bezeichnen wir $\alpha = n + 1$ als *Torsion* von k. Außerdem betrachten wir für einen auf E_2 liegenden Streckenzug von k den ersten Unterkreuzungspunkt im Sinne der Orientierung von k. Über ihm liegt ein Punkt eines Streckenzuges auf E_1, welcher der m-te Überkreuzungspunkt dieses Streckenzuges sei. Wir setzen $\beta = \alpha - m$, wenn die Unterkreuzung von rechts nach links erfolgt, $\beta = m - \alpha$ im anderen Falle und bezeichnen β als *Kreuzungsklasse* von k. Besitzt die Normalform keine Unterkreuzungen, so stellt sie entweder den Kreis oder zwei unverkettete Kreise dar. Im ersten Falle setzen wir $\alpha = \beta = 1$, im zweiten $\alpha = \beta = 0$. Durch Angabe von α und β ist die Normalform bis auf unwesentliche Deformationen bestimmt, so daß wir sie durch das Paar (α, β) bezeichnen können. Für $\alpha > 1$ ist β ungerade, teilerfremd zu α und $|\beta| < \alpha$, und zu jedem Paar (α, β), das diesen Bedingungen genügt, gehört eine Normalform. Bei ungeradem α stellt die Normalform einen einzigen Knoten dar, bei geradem α zwei verkettete Kreise.

Zwei Normalformen (α, β) und (α', β') stellen genau dann denselben Knoten bzw. dieselbe Verkettung dar, wenn $\alpha = \alpha'$ und $\beta = \beta'$ oder $\beta\beta' \equiv 1$ modulo 2α

[1]) Dies folgt unmittelbar aus der Definition der Viergeflechte und SCHUBERT [6] § 2. Über Viergeflechte s. BANKWITZ und SCHUMANN [1] und REIDEMEISTER [4].

ist (§ 3). Die hier betrachteten Knoten und Verkettungen sind also durch die Brückenzahl 2, die Torsion und die beiden möglichen Werte der Kreuzungsklasse vollständig charakterisiert. Sie sind sämtlich symmetrisch (§ 3). Da aus der Normalform (α, β) durch Spiegelung an einer senkrechten Ebene die Normalform $(\alpha, -\beta)$ entsteht, ist ein Knoten zur Normalform (α, β) genau dann amphicheiral, wenn $\beta^2 \equiv -1$ modulo 2α ist, und nur die Normalform (0,0) stellt eine amphicheirale Verkettung dar.

Zu den betrachteten Knoten gehören die Torusknoten der Umlaufzahl 2. Für diese ist $\beta = \pm 1$ und $\alpha\beta$ die Verschlingungszahl. Ferner gehören hierher diejenigen Schlingknoten, deren Diagonalknoten der Kreis ist (§ 3). Diese besitzen Normalformen der Gestalt $(\alpha, \pm(\alpha - 2))$, womit sich eine Vermutung von H. SEIFERT [9] bestätigt, daß der Kreis und der Viererknoten die einzigen amphicheiralen Schlingknoten sind.

Die hier betrachteten Knoten und Verkettungen stehen in einer engen Beziehung zu den Linsenräumen, die von H. SEIFERT bemerkt wurde. Sei k eine Brückendarstellung mit zwei Brücken. Wird der R^3 zur 3-Sphäre S^3 geschlossen, so ist die zweiblättrige, längs k verzweigte Überlagerung der S^3 ein Linsenraum, und es können alle Linsenräume so dargestellt werden (§ 4). Damit läßt sich die angegebene Klassifikation erhalten bis auf diejenigen Verkettungen, deren Verschlingungszahl null ist. Um auch diese mit zu erfassen, wird der Klassifikationssatz mit kombinatorischen Methoden bewiesen (§§ 5 bis 10).

Im Anhang geben wir Normalformen für diejenigen Knoten der ALEXANDER-BRIGGSschen Tabelle an, für die Brückendarstellungen mit zwei Brücken gefunden wurden.

§ 1. Brückendarstellungen

Es sei $\mathfrak{E}^3$ ein 3-Simplex, das der euklidische Raum R^3 mit einer 3-Sphäre S^3 gemein hat, die Rand eines euklidischen 4-Simplexes ist. R^3 und S^3 seien mit einer festen Orientierung versehen und zwar so, daß die Orientierungen auf $\mathfrak{E}^3$ übereinstimmen. Alle zu betrachtenden Knotenlinien und Flächen sollen in $\mathfrak{E}^3$ liegen, was weder für die Einbettung in den R^3 noch für die Einbettung in die S^3 eine Einschränkung der Allgemeinheit bedeutet[2]). Werden semilineare Abbildungen (s-Abbildungen) des R^3 oder der S^3 ausgeführt[3]), so sollen auch die Bilder der jeweils betrachteten Knotenlinien und Flächen in $\mathfrak{E}^3$ liegen. Die s-Abbildungen des R^3 sind dabei immer noch so abgeändert zu denken, daß sie außerhalb $\mathfrak{E}^3$ die Identität sind und somit eine s-Abbildung der S^3 induzieren. Alle Komplexe sollen stets eine Zerlegung in euklidische Simplexe besitzen.

Sei $\mathfrak{K}$ eine Kugel[4]). $k_1, k_2, \ldots, k_n$ seien Sehnen von $\mathfrak{K}$, d. h. paarweise disjunkte, orientierte Streckenzüge auf $\mathfrak{K}$, die mit dem Rande von $\mathfrak{K}$ nur die

[2]) Vgl. hierzu [6] § 1.

[3]) Über s-Abbildungen s. GRAEUB [3]. Die benutzten Resultate finden sich auch in [7] zusammengestellt. Alle s-Abbildungen sollen orientierungserhaltend sein.

[4]) Wir verstehen hierunter das semilineare Bild eines 3-Simplexes.

9272-65 10 R

Endpunkte gemein haben. Wenn es n paarweise disjunkte Elementarflächenstücke $\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_n$ auf $\mathfrak{K}$ gibt mit der Eigenschaft, daß der Rand von $\mathfrak{e}_i$ $(i = 1, 2, \ldots, n)$ aus k_i und einem Streckenzug l_i auf dem Rande von $\mathfrak{K}$ besteht und daß l_i den Durchschnitt von $\mathfrak{e}_i$ mit dem Rande von $\mathfrak{K}$ ausmacht, so sagen wir, daß die *Sehnen* k_i *in* $\mathfrak{K}$ *einfach* sind oder daß eine *Kugel mit n einfachen Sehnen* vorliegt und daß die *Elementarflächenstücke* $\mathfrak{e}_i$ *in die Sehnen* k_i *von* $\mathfrak{K}$ *eingespannt sind.*

Wie wir unten zeigen, stehen Kugeln mit einfachen Sehnen in engem Zusammenhang mit *Brückendarstellungen*, bei denen wir durchweg die folgende Gestalt benutzen wollen: Es seien T_1 und T_2 zwei konzentrische Tetraeder, die bezüglich des gemeinsamen Mittelpunktes Z ähnlich liegen und von denen T_1 das innere bezüglich des R^3 sei. Eine *Brückendarstellung k mit n Brücken* besteht aus je n paarweise disjunkten Streckenzügen auf T_1 und auf T_2 und $2n$ Strecken auf Geraden durch Z, die sich zu einer oder mehreren orientierten Knotenlinien zusammensetzen. Dabei liegt k auf der von T_1 und T_2 berandeten Kugelschale, und wir bezeichnen T_1 als das *innere*, T_2 als das *äußere Grenztetraeder* von k. Es kann angenommen werden, daß k im Inneren des Projektionskegels von Z nach einer (2-dimensionalen) Seite von T_2 liegt, da sich dies stets durch eine *zentrische Abbildung* erreichen läßt, d. h. durch eine s-Abbildung des R^3, welche jedes zu T_1 konzentrische und ähnlich gelegene Tetraeder in sich überführt[5]).

Sei T ein solches Tetraeder zwischen T_1 und T_2. Durch T wird die S^3 in zwei Kugeln $\mathfrak{K}$ und $\mathfrak{L}$ zerlegt, und es bildet k in $\mathfrak{K}$ und $\mathfrak{L}$ je n Sehnen. Diese sind in $\mathfrak{K}$ bzw. $\mathfrak{L}$ einfach. Man erhält nämlich Elementarflächenstücke, die in diese Sehnen eingespannt sind, wenn man k von Z aus projiziert. Eine Brückendarstellung mit n Brücken gibt also Anlaß zu einer Zerlegung der S^3 in zwei Kugeln mit je n einfachen Sehnen. Davon gilt auch die Umkehrung:

Satz 1. *Die Verkettung* $\varkappa$ *von* $m \geq 1$ *Knoten werde durch ein System k von m Knotenlinien dargestellt.* $\varkappa$ *besitzt genau dann eine Brückendarstellung mit n Brücken, wenn sich die* S^3 *durch eine 2-Sphäre so in zwei Kugeln zerlegen läßt, daß k in jeder der beiden Kugeln n einfache Sehnen bildet.*

Beweis. Besitzt $\varkappa$ eine Brückendarstellung k' mit n Brücken, so gibt es nach dem Vorangehenden zu k' eine 2-Sphäre T' mit den genannten Eigenschaften. Eine s-Abbildung, die k' auf k abbildet, führt T' in eine 2-Sphäre der gewünschten Art über.

Sei umgekehrt T eine 2-Sphäre, welche die S^3 so in zwei Kugeln $\mathfrak{K}$, $\mathfrak{L}$ zerlegt, daß k in $\mathfrak{K}$ und $\mathfrak{L}$ je n einfache Sehnen bildet. In die Sehnen von $\mathfrak{K}$ bzw. $\mathfrak{L}$ spannen wir die Elementarflächenstücke $\mathfrak{e}_i$ bzw. $\mathfrak{f}_i$ $(i = 1, 2, \ldots, n)$ ein. Durch eine geeignete s-Abbildung wird T in ein Tetraeder übergeführt, wobei $\mathfrak{K}$ in das Innere von T (bezüglich des R^3) übergehe. Wird T auf seinen Mittelpunkt hinreichend wenig stetig zusammengezogen, so trifft das dabei aus T entstehende Tetraeder jedes der Elementarflächenstücke $\mathfrak{e}_i$ nur in einem einzigen Querschnitt. Es werde T so zu T_1 zusammengezogen. Entsprechend entstehe T_2 aus

[5]) Vgl. [6] § 2.

9272-65 10 R

T durch stetige Dehnung, wobei während der Dehnung das aus T entstehende Tetraeder jedes Elementarflächenstück f_i in nur einem Querschnitt treffe. T_1 bzw. T_2 schneidet von e_i bzw. f_i ein Elementarflächenstück im Inneren von T_1 bzw. im Äußeren von T_2 ab. Über dieses läßt sich das von k herrührende Randstück kombinatorisch isotop in den Schnitt von T_1 bzw. T_2 mit e_i bzw. f_i deformieren[6]). k ist danach eine verallgemeinerte Brückendarstellung 2. Art mit n Brücken im Sinne von [6] und läßt sich durch eine zentrische Abbildung in eine Brückendarstellung der oben beschriebenen Art überführen[7]).

Hilfssatz 1. *Es sei k ein System von Knotenlinien, dessen Durchschnitt mit der Kugel $\mathfrak{K}$ aus n Sehnen k_i ($i=1, 2, \ldots, n$) von $\mathfrak{K}$ bestehe. $\mathfrak{K}$ werde durch ein Elementarflächenstück e so in zwei Kugeln zerlegt, daß die Sehnen k_i auf e liegen und von e paarweise disjunkte Elementarflächenstücke abschneiden. $\mathfrak{K}'$ sei eine Kugel im Inneren von $\mathfrak{K}$, deren Durchschnitt mit e ein Elementarflächenstück e' ist, das mit jeder der Sehnen k_i genau eine Sehne k_i' von $\mathfrak{K}'$ gemein hat. Dann gibt es eine s-Abbildung, die $\mathfrak{K}'$ auf $\mathfrak{K}$, e' auf e und k auf sich abbildet.*

Beweis. Durch eine semilineare Hilfsabbildung kann erreicht werden, daß $\mathfrak{K}$ ein euklidischer Würfel ist, in dem e eben und parallel zu einer Würfelseite liegt[8]). Die Abbildung kann zudem so geschehen, daß eine der Sehnen, etwa k_i, eine Strecke auf e ist und daß k den Würfelrand in den Endpunkten von k_i geradlinig durchsetzt. Man erkennt nun, daß sich $\mathfrak{K}'+k$ so deformieren läßt, daß k in sich übergeht und k_i so auf $\mathfrak{K}'$ zu liegen kommt, daß die Endpunkte von k_i und nur diese gemeinsame Punkte der Ränder von $\mathfrak{K}$ und $\mathfrak{K}'$ werden, während der Durchschnitt von $\mathfrak{K}'$ mit e ein Elementarflächenstück bleibt und sich auf den von k_i verschiedenen Sehnen nicht ändert. Die entsprechende Konstruktion kann der Reihe nach für alle Sehnen von $\mathfrak{K}$ ausgeführt werden. Die Ränder von $\mathfrak{K}$ und $\mathfrak{K}'$ haben dann die Endpunkte der Sehnen k_i ($i=1, 2, \ldots, n$) gemein. Der Durchschnitt von $\mathfrak{K}'$ mit e ist ein Elementarflächenstück e', das die Sehnen k_i enthält und dessen Komplement bezüglich e aus n Elementarflächenstücken $f_1, f_2, \ldots, f_n$ besteht. Den Rand von f_i bilden ein Streckenzug s_i auf dem Rande von $\mathfrak{K}$ und ein Streckenzug s_i' auf dem Rande von $\mathfrak{K}'$, wobei f_i mit den Sehnen von $\mathfrak{K}$ nur die Endpunkte von s_i und s_i' gemein hat. Da sich s_i' über f_i kombinatorisch isotop in s_i deformieren läßt, kann man e' so in e überführen, daß k punktweise festbleibt und daß der Rand von e der Durchschnitt der Ränder von $\mathfrak{K}$ und $\mathfrak{K}'$ wird[9]). Das Komplement von $\mathfrak{K}'$ bezüglich $\mathfrak{K}$ besteht dann aus zwei Kugeln, deren Durchschnitt mit k nur aus den Sehnenendpunkten besteht und vermöge deren man den Rand von $\mathfrak{K}'$ in den Rand von $\mathfrak{K}$ bei festem k überführen kann.

---

[6]) GRAEUB [3], S. 33.

[7]) [6] § 2. Der dortige Hilfssatz 1 überträgt sich ohne weiteres auf den hier benötigten Sachverhalt, bei dem statt senkrechter Strecken zwischen waagerechten Ebenen zentral gerichtete Strecken zwischen T_1 und T_2 benötigt werden.

[8]) Vgl. hierzu die Konstruktion in [8] S. 42, 43.

[9]) Vgl. [7] S. 271.

§ 2. Kugeln mit zwei einfachen Sehnen

Für eine Kugel mit zwei einfachen Sehnen sind die beiden Elementarflächenstücke, die man in die beiden Sehnen einspannen kann, bis auf Deformation[10]) eindeutig bestimmt. Dies ergibt sich aus dem Beweis von

Satz 2. *k sei eine Knotenlinie, deren Durchschnitt mit der Kugel $\mathfrak{K}$ aus zwei einfachen Sehnen von $\mathfrak{K}$ bestehe. $\mathfrak{e}_1, \mathfrak{e}_2$ und $\mathfrak{e}_1', \mathfrak{e}_2'$ seien je zwei Elementarflächenstücke auf $\mathfrak{K}$, die in diese Sehnen eingespannt sind. Dann gibt es eine s-Abbildung, die $\mathfrak{K}$ auf sich und $\mathfrak{e}_1', \mathfrak{e}_2'$ auf $\mathfrak{e}_1, \mathfrak{e}_2$ abbildet und die auf k und außerhalb einer beliebig kleinen Umgebung von $\mathfrak{K}$ die Identität ist.*

Anmerkung. *Der Satz bleibt richtig, wenn k aus zwei Knotenlinien besteht. Ein entsprechender Satz gilt natürlich für Kugeln mit einer einfachen Sehne, dagegen nicht für mehr als zwei Sehnen.*

Wir beweisen zunächst:

Hilfssatz 2. *Auf einem Elementarflächenstück $\mathfrak{e}$ seien zwei mittlere Punkte A, B durch zwei doppelpunktfreie Streckenzüge v, w verbunden. Dann läßt sich v auf $\mathfrak{e}$ kombinatorisch isotop in w deformieren.*

Beweis. Durch kombinatorisch isotope Deformation von v läßt sich zunächst erreichen, daß der Durchschnitt von v und w nur aus den Endpunkten und isolierten Schnittpunkten besteht. Haben nun v und w nur die Endpunkte gemein, so beranden sie auf $\mathfrak{e}$ ein Elementarflächenstück, über welches v in w deformiert werden kann. Den Fall, daß v und w noch Schnittpunkte gemein haben, führen wir hierauf zurück durch folgende Reduktion:

Wir zählen die Endpunkte A und B als Schnittpunkte, solange noch mindestens ein echter Schnittpunkt vorhanden ist. Seien C und D zwei auf w aufeinanderfolgende Schnittpunkte. Sie schneiden aus w ein von weiteren Schnittpunkten freies Stück w' aus und aus v ein Stück v', das zusammen mit w' auf $\mathfrak{e}$ ein Elementarflächenstück $\mathfrak{g}$ berandet. Falls $\mathfrak{g}$ keinen der Punkte A, B im Inneren enthält, so macht v' den Durchschnitt von v und $\mathfrak{g}$ aus, und es kann mindestens ein Schnittpunkt von v und w dadurch beseitigt werden, daß man v' über $\mathfrak{g}$ in w' deformiert und anschließend auf das $\mathfrak{g}$ gegenüberliegende Ufer von w abhebt. Wenn A oder B Endpunkt von w' ist, so ist dieser Punkt beim Abheben festzulassen.

Die geschilderte Reduktion führen wir so oft wie möglich für auf w benachbarte Schnittpunkte aus. Der Durchschnitt von v und w besteht dann nur noch aus den Endpunkten A und B. Der Nachweis erfolgt indirekt. Seien $A = C_0, C_1, \ldots, C_n = B$ die verbliebenen Schnittpunkte von v und w in ihrer Reihenfolge auf w, und nehmen wir an, daß $n > 1$ ist. C_{i-1} und C_i $(i = 1, 2, \ldots, n)$ schneiden dann aus w ein Stück w_i und aus v ein Stück v_i aus. Das von C_i nach B führende Stück von v bezeichnen wir mit u_i. Nun beranden w_1 und v_1 auf $\mathfrak{e}$ ein Elementarflächenstück $\mathfrak{g}_1$, das nach dem Vorangehenden den Punkt B enthalten muß. u_1 liegt daher auf $\mathfrak{g}_1$ und auch w_2, weil sich v und w in C_1 durchsetzen. C_2 muß auf u_1 liegen (Fig. 1), da andernfalls w_2 von $\mathfrak{g}_1$ ein

[10]) im Sinne von [7] § 1.

von v_2 und w_2 berandetes Elementarflächenstück $\mathfrak{g}'$ abschnitte, das weder A noch B im Inneren enthielte, weil w_1 und damit u_1 nicht auf $\mathfrak{g}'$ liegen könnte. v_2 und w_2 beranden also auf $\mathfrak{g}_1$ ein Elementarflächenstück $\mathfrak{g}_2$, das mit dem Rande von $\mathfrak{g}_1$ nur den Punkt C_1 gemein hat und daher A nicht enthält. Also liegt B und damit u_2 auf $\mathfrak{g}_2$. Diese Schlußweise läßt sich unter jeweiliger Erhöhung der Indizes um 1 wiederholen. Man gelangt so zu einem von v_{n-1} und w_{n-1} berandeten Elementarflächenstück $\mathfrak{g}_{n-1}$, das den Punkt B, nicht aber A enthält und das auf $\mathfrak{g}_{n-2} \subset \ldots \subset \mathfrak{g}_1$ liegt. Dann liegen aber v_n und w_n auf $\mathfrak{g}_{n-1}$ und beranden auf $\mathfrak{g}_{n-1}$ ein Elementarflächenstück $\mathfrak{g}_n$, das A nicht enthält und B als Randpunkt besitzt, womit sich ein Widerspruch ergibt und die Behauptung folgt.

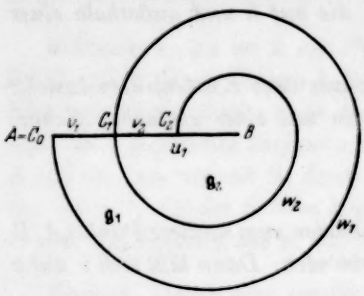


Fig. 1

Beweis von Satz 2. Es seien k_1, k_2 die beiden Sehnen und T der Rand von $\mathfrak{R}$. In k_1 seien $\mathfrak{e}_1$ und $\mathfrak{e}_1'$ eingespannt, und es seien w bzw. v die auf T liegenden Randstücke von $\mathfrak{e}_1$ bzw. $\mathfrak{e}_1'$. Wir führen zunächst $\mathfrak{e}_1'$ in $\mathfrak{e}_1$ über.

1. Schritt. *Reduktion der Schnitte von w und v.*

Die Reduktion entspricht derjenigen im Beweise von Hilfssatz 2, wobei hier verlangt wird, daß die dort mit v' und w' bezeichneten Stücke von v und w auf T ein Elementarflächenstück $\mathfrak{g}$ beranden, das keinen der Endpunkte von k_1 und k_2 im Inneren enthält. Sie werde so oft wie möglich ausgeführt. Jede Reduktion läßt sich auch durch eine s-Abbildung erhalten, die T in sich überführt und k punktweise festläßt.

2. Schritt. *Reduktion der Schnitte von $\mathfrak{e}_1$ und $\mathfrak{e}_1'$.*

Durch isotope simpliziale Deformation von $k + T + \mathfrak{e}_1'$, die T in sich überführt, k punktweise festläßt und die Zahl der Schnittpunkte von v und w nicht ändert, kann erreicht werden, daß der Durchschnitt von $\mathfrak{e}_1$ und $\mathfrak{e}_1'$ aus dem gemeinsamen Randstück k_1 und Schnittlinien besteht, die geschlossen oder möglicherweise Querschnitte sind und die höchstens Punkte von k_1 gemein haben. Querschnitte, deren beide Endpunkte auf k_1 liegen, denken wir uns durch das von den Endpunkten aus k_1 ausgeschnittene Stück geschlossen. Unter den geschlossenen Schnitten gibt es einen auf $\mathfrak{e}_1$ innersten s, der auf $\mathfrak{e}_1$ ein von weiteren Schnitten freies Elementarflächenstück $\mathfrak{f}$ berandet. Auf $\mathfrak{e}_1'$ berandet s ein Elementarflächenstück $\mathfrak{f}'$, wobei $\mathfrak{f}$ und $\mathfrak{f}'$ mit T höchstens die Endpunkte von k_1 gemein haben. $\mathfrak{f}$ und $\mathfrak{f}'$ beranden auf $\mathfrak{R}$ eine Kugel $\mathfrak{L}$. Gemeinsame Punkte von k und $\mathfrak{L}$ können nur auf s liegen und machen dann entweder nur einen Punkt oder ein zusammenhängendes Stück von k_1 aus. Durch eine s-Abbildung, die auf k und im Komplement von $\mathfrak{R}$ die Identität ist, kann der Schnitt s so beseitigt werden, daß keine neuen Schnitte entstehen. Es werden so der Reihe nach alle geschlossenen Schnitte von $\mathfrak{e}_1$ und $\mathfrak{e}_1'$ beseitigt. Danach sind alle Schnitte verschwunden:

3. Schritt. *Nachweis, daß der Durchschnitt von* $\mathfrak{e}_1$ *und* $\mathfrak{e}_1'$ *nur noch aus* k_1 *besteht.*

Nehmen wir an, daß noch Querschnitte von $\mathfrak{e}_1$ und $\mathfrak{e}_1'$ vorhanden sind. Unter diesen können solche sein, die einen Endpunkt auf k_1 besitzen, der nicht zugleich Endpunkt von k_1 ist. Wir denken uns diese durch ein Stück von k_1 so verlängert, daß beide Endpunkte auf T liegen. Jeder Querschnitt schneidet dann von $\mathfrak{e}_1$ ein Elementarflächenstück ab, das mit T genau ein zusammenhängendes Stück des Randes gemein hat. Sei $\mathfrak{f}$ ein auf $\mathfrak{e}_1$ innerstes solches Elementarflächenstück. Es wird berandet von einem Querschnitt s und einem Stück t von w, und es enthält $\mathfrak{f}$ außer dem Randbogen s keinen weiteren Querschnitt. Von $\mathfrak{e}_1'$ schneidet s ein Elementarflächenstück $\mathfrak{f}'$ ab, das von s und einem Stück t' von v berandet wird, wobei t und t' genau die Endpunkte gemein haben. $\mathfrak{f}+\mathfrak{f}'$ ist ein Elementarflächenstück, das $\mathfrak{K}$ in zwei Kugeln zerlegt. Eine von

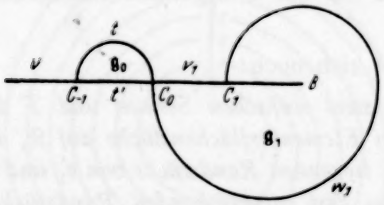


Fig. 2a

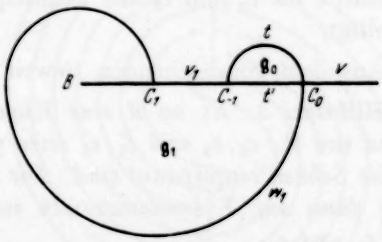


Fig. 2b

beiden, sie heiße $\mathfrak{L}$, enthält keinen Punkt von k_1 im Inneren. Der Durchschnitt von $\mathfrak{L}$ und T ist ein Elementarflächenstück $\mathfrak{g}_0$, das von t und t' berandet wird und keinen Endpunkt von k_1 im Inneren enthält. Der Durchschnitt von $\mathfrak{g}_0$ und v besteht daher aus $t' \cdot k_2$ liegt auf $\mathfrak{L}$, weil andernfalls $\mathfrak{g}_0$ keinen der Endpunkte von k_1 und k_2 im Inneren enthielte im Widerspruch zum 1. Schritt.

Seien C_{-1} und C_0 die Endpunkte von t. Mindestens einer von beiden, etwa C_0, ist nicht Endpunkt von k_1. C_0 zerlegt w in zwei Stücke. Dabei sei w_0 dasjenige, das mit t nur den Endpunkt C_0 gemein hat. Wir denken uns w_0 in C_0 entspringend und betrachten den auf w_0 nächsten Schnittpunkt C_1 von v und w. C_0 und C_1 schneiden aus w ein Stück w_1 und aus v ein Stück v_1 aus. w_1 hat mit $\mathfrak{g}_0$ höchstens die Endpunkte gemein, weil sich v und w in C_0 schneiden und der Durchschnitt von $\mathfrak{g}_0$ und v aus t' besteht, und es beranden v_1 und w_1 auf T ein Elementarflächenstück $\mathfrak{g}_1$, das mit $\mathfrak{g}_0$ keine inneren Punkte gemein hat. $\mathfrak{g}_1$ enthält mindestens einen der Endpunkte von k_1 im Inneren, da sonst ein Widerspruch zum 1. Schritt vorläge. Es kann auch nur ein Endpunkt von k_1 im Inneren von $\mathfrak{g}_1$ liegen (Fig. 2a und b). Da sich nämlich v und w in C_0 durchsetzen und da $\mathfrak{g}_0$ und $\mathfrak{g}_1$ in der Umgebung von C_0 an verschiedene Ufer von v anstoßen, kann nur eines der beiden Stücke, in welches v durch C_0 zerlegt wird, auf $\mathfrak{g}_1$ liegen, während das andere mit $\mathfrak{g}_1$ nur den Endpunkt C_0 gemein hat. Sei B der auf $\mathfrak{g}_1$ liegende Endpunkt von k_1 und v_0 das von C_0 nach B führende Stück von v. Es kann nun die Schlußweise des Beweises von Hilfssatz 2 auf w_0 und v_0 angewandt werden, womit sich der gesuchte Widerspruch ergibt.

4. Schritt. *Abschluß des Beweises.*

Nach dem Vorangehenden bilden e_1 und e_1' zusammen ein Elementarflächenstück, welches von $\mathfrak{K}$ eine Kugel $\mathfrak{L}$ abschneidet, die k_2 nicht enthält. Der Durchschnitt von $\mathfrak{L}$ und T ist ein von v und w berandetes Elementarflächenstück, über welches v kombinatorisch isotop in w deformiert werden kann. Da man annehmen kann, daß T ein Tetraeder und e_1 eben ist, läßt sich das Deformationsresultat durch eine s-Abbildung erhalten, die k punktweise festläßt, T auf sich abbildet und e_1' so, daß der Durchschnitt von e_1 und e_1' aus dem gemeinsamen Rande besteht[11]). Danach läßt sich e_1' bei festem k und T in e_1 überführen.

e_2' wird nun bei festem $e_1' = e_1$ entsprechend in e_2 übergeführt. Der 3. Schritt ist dabei überflüssig, da man vermöge Hilfssatz 2 schließen kann, daß der 1. Schritt für e_2 und e_2' alle Schnittpunkte der auf T liegenden Randstücke beseitigt.

Aus dem vorangehenden Beweis ergibt sich noch:

Hilfssatz 3. *Es sei $\mathfrak{K}$ eine Kugel mit zwei einfachen Sehnen und T der Rand von $\mathfrak{K}$. e_1, e_2 und e_1', e_2' seien je zwei Elementarflächenstücke auf $\mathfrak{K}$, die in die Sehnen eingespannt sind. Die auf T liegenden Randstücke von e_1' und e_2' sind dann auf T kombinatorisch isotop zu den entsprechenden Randstücken von e_1 und e_2.*

§ 3. Normalformen

Es sei k eine Brückendarstellung im Sinne von § 1 mit zwei Brücken, T_1 das innere, T_2 das äußere Grenztetraeder und Z ihr Mittelpunkt. Die beiden auf T_1 bzw. T_2 liegenden Stücke von k bezeichnen wir als *innere* bzw. *äußere Wege*. Eine Strecke heiße *zentral*, wenn sie auf einer Geraden durch Z liegt. k besteht also aus den beiden inneren Wegen, den beiden äußeren Wegen und vier zentralen Strecken zwischen T_1 und T_2. Eine *Brücke* von k wird von einem inneren Weg und den beiden anstoßenden Strecken gebildet. Projiziert man k von Z aus auf T_2, so bildet sich jede zentrale Strecke in einen Punkt ab, während jeder andere Doppelpunkt der Projektion Bild von genau zwei Punkten auf k ist, nämlich von einem *Überkreuzungspunkt* auf einem inneren Weg und einem *Unterkreuzungspunkt* auf einem äußeren Weg.

Wir normieren die Brückendarstellung dadurch, daß wir Unterkreuzungen so weit wie möglich beseitigen. Seien w_1, w_2 die Projektionen der inneren Wege, v_1, v_2 die äußeren Wege. Zunächst können v_1 und v_2 bei festen Endpunkten auf T_2 so deformiert werden, daß v_1 und v_2 mit w_1 und w_2 nur die Endpunkte und Schnittpunkte gemein haben. Die Endpunkte von v_1 und v_2 zählen als Schnittpunkte, solange auf v_1 bzw. v_2 noch mindestens ein weiterer Schnittpunkt liegt.

Nehmen wir an, daß einer der beiden Wege w_1, w_2, etwa w_1, zweimal hintereinander denselben äußeren Weg, etwa v_2, schneidet. Die beiden

[11]) Vgl. [7] S. 271.

Schnittpunkte schneiden aus w_1 ein von weiteren Schnitten freies Stück w', aus v_2 ein Stück v' aus. Es gibt dann ein Elementarflächenstück $\mathfrak{g}$, das von w' und v' berandet wird und das, abgesehen von seinem Rande, außerhalb T_2 liegt. Über $\mathfrak{g}$ läßt sich v' kombinatorisch isotop in w' deformieren und danach so von w' abheben, daß mindestens ein Schnittpunkt verschwunden ist. Die Endpunkte von v_1 und v_2 sind dabei festzulassen. Die angegebene Reduktion werde so oft wie möglich ausgeführt. Danach ist eine Normalform von k erreicht: Entweder haben w_1, w_2 mit v_1, v_2 nur noch die Endpunkte gemein oder jeder der beiden Wege w_1 und w_2 trifft abwechselnd v_1 und v_2. Im

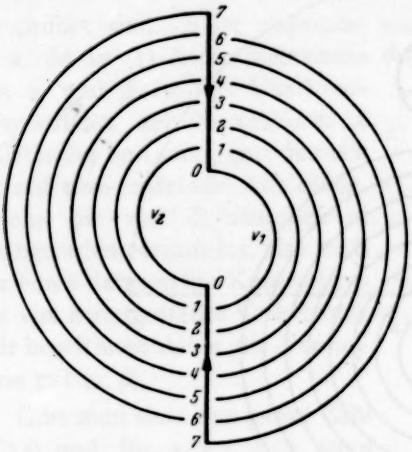


Fig. 3

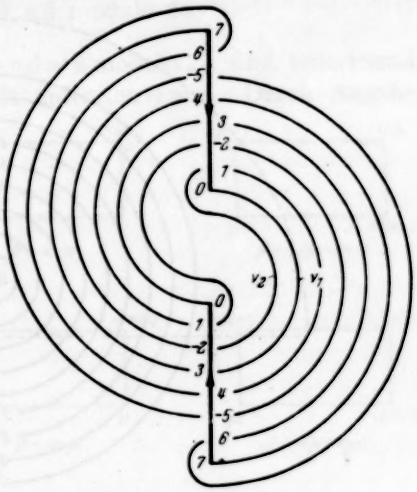


Fig. 4

zweiten Falle trifft auch jeder der beiden Wege v_1 und v_2 abwechselnd w_1 und w_2: Träfe etwa v_1 zweimal hintereinander w_2, so berandeten diese beiden Schnittpunkte auf v_1 ein von weiteren Schnitten freies Stück, während sie auf w_2 ein Stück mit einer ungeraden Anzahl von Schnittpunkten von w_2 mit v_1, v_2 berandeten, weil w_2 abwechselnd v_1 und v_2 trifft. Dies ist aber unmöglich, weil sich w_1, w_2, v_1, v_2 zu einem oder zwei geschlossenen Wegen (mit Doppelpunkten) zusammensetzen.

Besitzt die *Normalform* keine Überkreuzungen, so stellt sie entweder den Kreis oder zwei unverkettete Kreise dar. Im ersten Falle bezeichnen wir sie mit (1, 1), im zweiten Falle mit (0, 0).

Besitzt die *Normalform* k Überkreuzungen, so rechnen wir die Endpunkte der inneren Wege darunter. Jeder innere Weg überkreuzt die äußeren Wege abwechselnd, und jeder äußere Weg unterkreuzt die inneren Wege abwechselnd. Die Anzahl der Überkreuzungen ist von der Form $2n$ mit $n > 2$. Wir bezeichnen $\alpha = n - 1$ als *Torsion*[12]) von k. Fig. 3–5 zeigen Projektionen von Normalformen für $\alpha = 7$. Ist α gerade, so sind die Endpunkte eines äußeren

[12]) Die Abweichung gegenüber der Einleitung beruht darauf, daß hier die Endpunkte der inneren Wege als Überkreuzungen zählen.

Weges von k Endpunkte derselben Brücke, weil die äußeren Wege die beiden inneren abwechselnd unterkreuzen, und es stellt k zwei verkettete Kreise dar. Ist α ungerade, so stellt k einen einzigen Knoten dar.

Seien u_1, u_2 die inneren, v_1, v_2 die äußeren Wege der Normalform k, wobei v_1 unter dem Endpunkt von u_1 entspringe. Für $\alpha > 1$ bezeichnen wir die Überkreuzungspunkte auf u_1 und auf u_2 mit $0, 1, \ldots, \alpha$, und zwar in der Reihenfolge, in der sie entgegen der Orientierung von k auftreten. Den Unterkreuzungspunkten ordnen wir Restklassen modulo 2α zu, die wir durch die ganzen Zahlen des halboffenen Intervalls $(-\alpha, \alpha]$ repräsentieren: Der

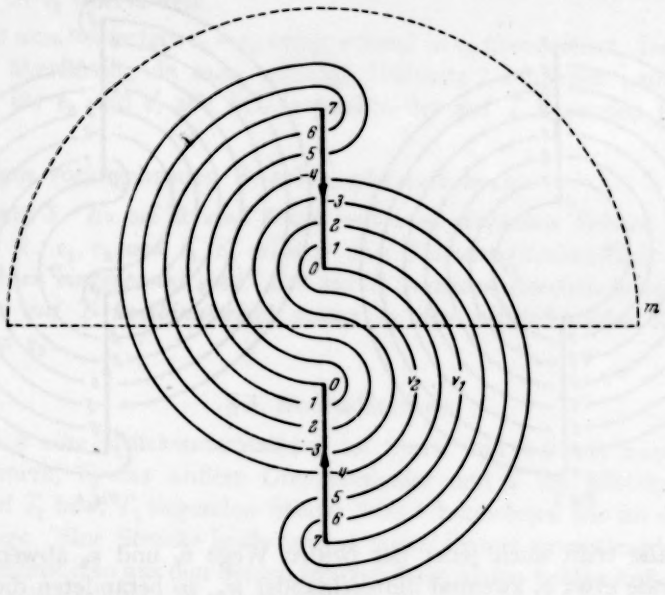


Fig. 5

Unterkreuzungspunkt unter dem Überkreuzungspunkt γ erhält die Restklasse γ zugeordnet, wenn die Unterkreuzung von rechts nach links erfolgt, $-\gamma$ im anderen Falle. Diese Zuordnung ist auch in den Endpunkten von v_1 und v_2 wegen $\gamma = 0$ bzw. $\gamma = \alpha$ sinnvoll (Fig. 3–5). Durch die Unterkreuzungspunkte zerfallen v_1 und v_2 in je α Stücke. Für diese subtrahieren wir die Restklasse des Anfangspunktes von derjenigen des Endpunktes. Die Differenz β ist eine von der Auswahl des Stückes unabhängige Restklasse: v_1, v_2 und die Projektionen w_1, w_2 von u_1, u_2 zerlegen das äußere Grenztetraeder T_2 in 2α Elementarflächenstücke, wobei jedes von je einem Stück von v_1, v_2, w_1, w_2 berandet wird. Sei e ein solches Elementarflächenstück, dessen Ecken den Überkreuzungen der Nummern $\mu, \mu+1$ und $\nu, \nu+1$ entsprechen mögen. Fig. 6a–f zeigt für alle möglichen Orientierungen der Randstücke von e, daß den zu v_1 und v_2 gehörigen Randstücken dieselbe Differenzklasse zugeordnet ist. Dabei sind die zu w_1, w_2 gehörigen Randstücke senkrecht, die zu v_1, v_2 gehörigen Randstücke waagerecht gezeichnet,

und an jeder Ecke ist die zur Unterkreuzung gehörige Restklasse angeschrieben. Da jedes der Elementarflächenstücke, in welche T_2 durch v_1, v_2, w_1 und w_2 zerlegt wird, die Eigenschaft hat, daß seinen beiden zu v_1, v_2 gehörigen Randstücken dieselbe Restklasse zugeordnet wird, ist die Restklasse β für alle Stücke von v_1 und v_2 dieselbe.

Wir bezeichnen β als die *Kreuzungsklasse* der Normalform k. Durchläuft man einen äußeren Weg, etwa v_1, so gelangt man der Reihe nach zu Unterkreuzungen, die abwechselnd unter u_1 und u_2 liegen und denen die Restklassen

$$0, \beta, 2\beta, \ldots, (\alpha-1)\beta, \alpha\beta \text{ modulo } 2\alpha \tag{1}$$

zugeordnet sind. β ist ungerade wegen $\alpha\beta \equiv \alpha$ modulo 2α und teilerfremd zu α, da in (1) keine Restklasse doppelt auftreten kann. Durch Angabe von α und β ist das Überkreuzungsschema der Normalform k vollständig bestimmt und damit k bis auf eine zentrische Abbildung, welche die von Z ausgehenden Halbgeraden permutiert, also auch der von k dargestellte Knoten bzw. die von k dargestellte Verkettung. Wir bezeichnen daher die *Normalform* mit (α, β).

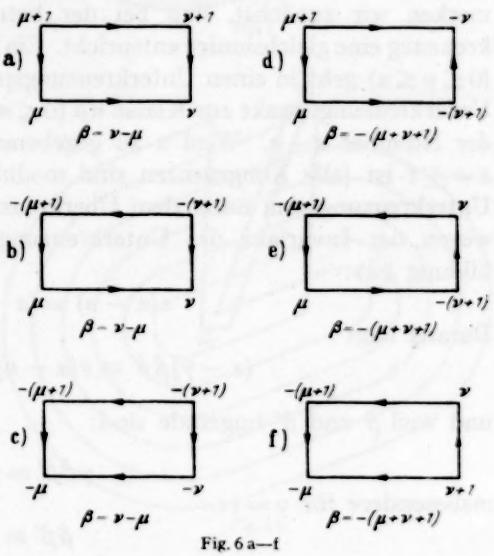


Fig. 6 a—f

Gibt man sich eine ganze Zahl $\alpha \geq 0$ und für $\alpha > 1$ eine zu α teilerfremde, ungerade ganze Zahl β im Intervall $(-\alpha, \alpha)$ vor, so gehört dazu offenbar auch eine Normalform.

Für spätere Anwendung bemerken wir: Werden bei einer Normalform (α, β) mit $\alpha \geq 1$ auf dem äußeren Grenztetraeder δ geschlossene, paarweise punktfremde Wege gezeichnet, so daß jeder dieser Wege die äußeren Wege der Normalform trennt, so kann dies so geschehen, daß die geschlossenen Wege die inneren Wege der Normalform abwechselnd unterkreuzen, und zwar je α-mal. Zählt man die so erhaltenen Unterkreuzungen mit, so liest man an Stelle von α und β nun $\bar{\alpha} = \alpha(\delta+1)$ und $\bar{\beta} = \beta(\delta+1)$ ab, und es ist $\delta+1$ der größte gemeinsame Teiler von $\bar{\alpha}$ und $\bar{\beta}$.

Orientiert man eine Normalform (α, β) um, wobei im Falle einer Verkettung beide Komponenten umzuorientieren sind, so bleiben α und die Unterkreuzungssinne erhalten. Eine bisher mit γ bezeichnete Überkreuzung wird nun mit $\alpha - \gamma$ bezeichnet, und für jedes Stück der äußeren Wege vertauschen sich Anfangs- und Endpunkt, woraus folgt, daß auch β erhalten bleibt. Bezeichnet man eine Verkettung als symmetrisch, wenn sie bei Umorientierung aller Komponenten in sich übergeht, so gilt also

Satz 3. *Knoten und Verkettungen, die Brückendarstellungen mit zwei Brücken besitzen, sind symmetrisch*[13]).

Stellt k eine Verkettung dar und orientiert man nur eine der Komponenten um, so geht β in die Klasse $\alpha+\beta$ modulo 2α über.

Wir hatten bemerkt, daß ein Paar (α, β), das den oben angegebenen Bedingungen genügt, eindeutig einen Knoten bzw. eine Verkettung bestimmt. Das Umgekehrte gilt im allgemeinen nicht. Dies beruht darauf, daß bei der Definition der Kreuzungsklasse die äußeren vor den inneren Wegen ausgezeichnet sind. Es gibt offenbar eine (orientierungserhaltende) s-Abbildung der S^3, welche die Grenztetraeder T_1, T_2 vertauscht und jede zentrale Strecke zwischen T_1 und T_2 wieder in eine solche überführt. Die Normalform (α, β) geht dabei in eine Normalform (α, β') über. Zur Bestimmung von β' bemerken wir zunächst, daß bei der betrachteten Abbildung jeder Unterkreuzung eine gleichsinnige entspricht. Ein Überkreuzungspunkt der Nummer γ $(0 \leq \gamma \leq \alpha)$ geht in einen Unterkreuzungspunkt zur Klasse $(\alpha-\gamma)\beta'$ über, ein Unterkreuzungspunkt zur Klasse $n\beta$ $(0 \leq n \leq \alpha)$ in einen Überkreuzungspunkt der Nummer $\alpha-n$. Wird n zu gegebenem γ so gewählt, daß $n\beta \equiv \varepsilon\gamma$ mit $\varepsilon = \pm 1$ ist (alle Kongruenzen sind modulo 2α zu verstehen), so gibt ε den Unterkreuzungssinn unter dem Überkreuzungspunkt der Nummer γ an, und wegen der Invarianz des Unterkreuzungssinnes bei der betrachteten Abbildung gilt:

$$\varepsilon(\alpha - n) \equiv (\alpha - \gamma)\beta'.$$

Daraus folgt

$$(\alpha - \gamma)\beta\beta' \equiv \varepsilon(\alpha - n)\beta \equiv \varepsilon\alpha\beta - \gamma,$$

und weil β und β' ungerade sind:

$$\gamma\beta\beta' \equiv \gamma,$$

insbesondere für $\gamma = 1$:

$$\beta\beta' \equiv 1.$$

Zwei Normalformen (α, β) und (α, β') stellen also denselben Knoten bzw. dieselbe Verkettung dar, wenn $\beta\beta' \equiv 1$ modulo 2α ist. Dies ist beispielsweise für Fig. 4 und Fig. 5 der Fall. Damit sind aber alle Darstellungsmöglichkeiten erschöpft. Es gilt:

Satz 4. *Zwei Normalformen (α, β) und (α', β') stellen genau dann denselben Knoten bzw. dieselbe Verkettung dar, wenn $\alpha = \alpha'$ und $\beta = \beta'$ oder $\beta\beta' \equiv 1$ modulo 2α ist.*

Den Beweis dieses Klassifikationssatzes erbringen wir in den §§ 5–10.

Durch Spiegelung an einer Ebene im R^3 erhält man aus der Normalform (α, β) die Normalform $(\alpha, -\beta)$, weil sich die Unterkreuzungssinne umkehren. Daher gilt:

[13]) Für Knoten wurde dies von BANKWITZ und SCHUMANN [1] auf anderem Wege bewiesen.

Korollar zu Satz 4. *Ein durch die Normalform* (α, β) *darstellbarer Knoten ist genau dann amphicheiral, wenn* $\beta^2 \equiv -1$ *modulo* 2α *ist. Keine von* $(0, 0)$ *verschiedene Normalform stellt eine amphicheirale Verkettung dar.*

Für ungerades $\alpha > 1$ und $\beta = \pm 1$ stellt die Normalform (α, β) offenbar einen Torusknoten der Umlaufzahl 2 und der Verschlingungszahl $\alpha\beta$ dar (vgl. Fig. 3). Wegen Satz 4 sind damit alle in Frage kommenden Torusknoten erfaßt, da unter den Torusknoten nur diejenigen der Umlaufzahl 2 eine Brückendarstellung mit zwei Brücken besitzen[14]).

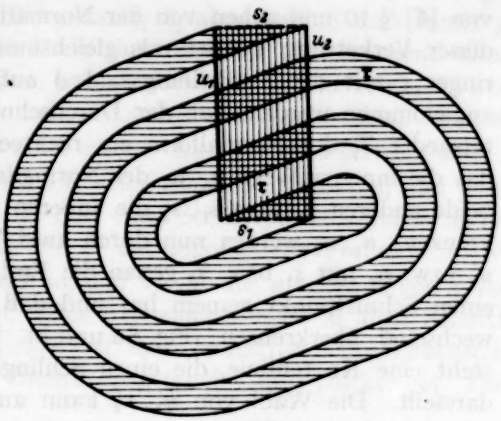


Fig. 7

Satz 5. *Jeder Schlingknoten, der den Kreis als Diagonalknoten besitzt, läßt sich durch eine Normalform der Gestalt* $(\alpha, \pm(\alpha - 2))$ *darstellen. Kreis und Viererknoten sind die einzigen amphicheiralen Schlingknoten*[15]).

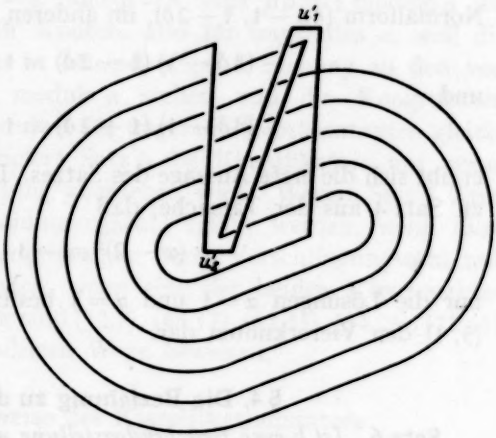


Fig. 8a

Beweis. Schlingknoten, deren Diagonalknoten vom Kreis verschieden ist, besitzen keine Brükkendarstellung mit zwei Brücken (SCHUBERT [6]) und sind nicht amphicheiral (SEIFERT [9]). Der Kreis wird durch die Normalform $(1, 1)$ dargestellt und kann daher im Folgenden ausgeschlossen werden. Wir können annehmen, daß die Verdrillungszahl δ des Schlingknotens positiv ist, da sich negative Verdrillungszahlen durch Spiegelung erhalten lassen. Wir benutzen die Konstruktion

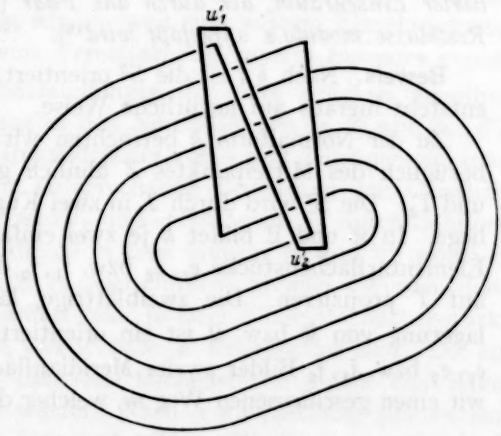


Fig. 8b

[14]) [6] Satz 10.

[15]) Über Schlingknoten s. SEIFERT [9], GRAEUB [3], SCHUBERT [6], [7].

von [6] § 10 und gehen von der Normalform $(2\delta, 1)$ aus. Die beiden Kreise dieser Verkettung können als gleichsinnig orientierte Ränder eines Kreisringes $\mathfrak{r}$ mit der Verdrillungszahl δ aufgefaßt werden (Fig. 7). Es kann angenommen werden, daß der Durchschnitt von $\mathfrak{r}$ mit dem inneren Grenztetraeder T_1 der Normalform ein rechteckiges Elementarflächenstück $\mathfrak{e}$ ist, das die inneren Wege u_1, u_2 der Normalform als Kanten besitzt und dessen beide anderen Kanten s_1, s_2 die äußeren Wege der Normalform nicht überkreuzen. u_1, u_2 werden nun durch zwei Wege u_1', u_2' auf T_1 so ersetzt, daß u_1' bzw. u_2' mit s_1 bzw. s_2 genau die Endpunkte und mit s_2 bzw. s_1 genau einen Schnittpunkt gemein hat und daß u_1' und u_2' die äußeren Wege abwechselnd überkreuzen (Fig. 8a und b). Durch kohärente Orientierung entsteht eine Knotenlinie, die einen Schlingknoten mit der Verdrillungszahl δ darstellt. Die Wahl von u_1', u_2' kann auf zwei zueinander spiegelbildliche Weisen erfolgen, die beiden zugehörigen Schlingknoten unterscheiden sich durch das Vorzeichen der Eigenschnittzahl. In dem einen Falle erhält man die Normalform $(4\delta - 1, 1 - 2\delta)$, im anderen $(4\delta + 1, 1 + 2\delta)$. Wegen Satz 4 und

$$-(4\delta - 3)(1 - 2\delta) \equiv 1 \mod (8\delta - 2)$$

und

$$-(4\delta - 1)(1 + 2\delta) \equiv 1 \mod (8\delta + 2)$$

ergibt sich die erste Aussage des Satzes. Die zweite folgt wegen des Korollars zu Satz 4 aus der Tatsache, daß

$$(\alpha - 2)^2 \equiv -1 \mod 2\alpha$$

nur die Lösungen $\alpha = 1$ und $\alpha = 5$ besitzt, denn es stellt $(1, 1)$ den Kreis, $(5, 3)$ den Viererknoten dar.

§ 4. Die Beziehung zu den Linsenräumen

Satz 6. *Ist k eine Brückendarstellung mit zwei Brücken von der Normalform (α, β), so ist die zweiblättrige, längs k verzweigte Überlagerung der S^3 ein orientierter Linsenraum, der durch das Paar (α, β) charakterisiert ist, wenn β als Restklasse modulo α aufgefaßt wird*[16]).

Beweis. Nach § 1 ist die S^3 orientiert, die Orientierung der Überlagerung entsteht hieraus auf natürliche Weise.

Zu der Normalform k betrachten wir ein zu den Grenztetraedern T_1, T_2 bezüglich des Mittelpunktes Z ähnlich gelegenes Tetraeder T zwischen T_1 und T_2. Die S^3 wird durch T in zwei Kugeln $\mathfrak{K}$ und $\mathfrak{L}$ zerlegt, wobei Z auf $\mathfrak{K}$ liege. In $\mathfrak{K}$ und $\mathfrak{L}$ bildet k je zwei einfache Sehnen. In diese spannen wir Elementarflächenstücke $\mathfrak{e}_1, \mathfrak{e}_2$ bzw. $\mathfrak{f}_1, \mathfrak{f}_2$ dadurch ein, daß wir sie von Z aus auf T projizieren. Die zweiblättrige, längs der Sehnen verzweigte Überlagerung von $\mathfrak{K}$ bzw. $\mathfrak{L}$ ist ein orientierter Vollring $\hat{\mathfrak{K}}$ bzw. $\hat{\mathfrak{L}}$, und es sind $\mathfrak{e}_1, \mathfrak{e}_2$ bzw. $\mathfrak{f}_1, \mathfrak{f}_2$ Bilder zweier Meridianflächen von $\hat{\mathfrak{K}}$ bzw. $\hat{\mathfrak{L}}$. Auf T wählen wir einen geschlossenen Weg m, welcher die auf T liegenden Randstücke von

[16]) Dieser Satz stammt von H. SEIFERT.

$\mathfrak{e}_1$ und $\mathfrak{e}_2$ trennt und die Randstücke von $\mathfrak{f}_1, \mathfrak{f}_2$ abwechselnd und damit je α-mal trifft. In Fig. 5 ist m gestrichelt eingezeichnet. Entsprechend kann auf T ein Weg n gezeichnet werden, so daß dabei die Rollen von $\mathfrak{e}_1, \mathfrak{e}_2$ und $\mathfrak{f}_1, \mathfrak{f}_2$ vertauscht sind und sich m und n in 2α Punkten schneiden. Über m liegen zwei Meridiane von $\hat{\mathfrak{R}}$, sei $\hat{m}$ einer davon. Entsprechend sei $\hat{n}$ ein über n liegender Meridian von $\hat{\mathfrak{L}}$. $\hat{m}$ und $\hat{n}$ schneiden sich in α Punkten, und jeder zweite Schnittpunkt von m und n ist Bild eines Schnittpunktes von $\hat{m}$ und $\hat{n}$. Ist $\alpha = 0$, so ist die Überlagerung der S^3 das topologische Produkt $S^1 \times S^2$. Ist $\alpha > 0$, so orientieren wir $\hat{m}$ und $\hat{n}$ so, daß von $\hat{\mathfrak{R}}$ aus gesehen $\hat{m}$ von $\hat{n}$ von links nach rechts durchsetzt wird. (Der Kreuzungssinn ist in allen Schnittpunkten derselbe.) Geht man nun auf $\hat{n}$ von einem Schnittpunkt zum nächsten, so rückt man auf $\hat{m}$ im Sinne der Orientierung um β Schnittpunkte vor. Dies ist die Behauptung.

Nach REIDEMEISTER [5] ergeben zwei Paare (α, β) und (α', β') denselben orientierten Linsenraum, wenn $\alpha = \alpha'$ und $\beta = \beta'$ oder $\beta\beta' \equiv 1$ modulo α ist. Damit folgt Satz 4 aus Satz 6 für Knoten, also für ungerades α, weil die ungeraden Restklassen modulo 2α in eineindeutiger Beziehung zu den von null verschiedenen Restklassen modulo α stehen und die Kongruenzen $\beta\beta' \equiv 1$ modulo 2α und $\beta\beta' \equiv 1$ modulo α für ungerade Repräsentanten gleichwertig sind. Für Verkettungen liefert Satz 6 die Klassifikation nur, wenn von der Orientierung der Verkettung abgesehen wird. Wegen Satz 3 kann die Orientierung durch die Verschlingungszahl erfaßt werden, wenn diese nicht null ist. Es gibt jedoch Verkettungen, deren Verschlingungszahl null ist und die nicht in sich übergehen, wenn eine der beiden Komponenten umorientiert wird. Beispiele hierfür sind (32,7), (48,17), (64,15), (64,23), (72,11). Wir müssen daher Satz 4 auf anderem Wege beweisen.

§ 5. Ansatz zum Beweise des Klassifikationssatzes

Um zusätzliche Komplikationen zu vermeiden, schließen wir die Normalformen (0, 0) und (1, 1) im Folgenden aus. Dies ist wegen Satz 6 möglich. Alle anderen Normalformen lassen wir zu, weil eine weitere Einschränkung, die nach Satz 6 möglich wäre, keine Vereinfachung unseres Beweises ergibt.

Seien k und k' zwei Brückendarstellungen mit zwei Brücken desselben Knotens bzw. derselben Verkettung in den Normalformen (α, β) und (α', β'). T_1 sei das innere, T_2 das äußere Grenztetraeder von k und Z ihr gemeinsamer Mittelpunkt. T_1', T_2', Z' sollen die entsprechende Bedeutung für k' haben. Das Tetraeder S' liege zwischen T_1' und T_2' und entstehe aus T_1' durch Dehnung mit dem Zentrum Z'. S' zerlegt die S^3 in zwei Kugeln $\mathfrak{K}'$ und $\mathfrak{L}'$, wobei Z' auf $\mathfrak{K}'$ liege. k' bildet in $\mathfrak{K}'$ und $\mathfrak{L}'$ je zwei einfache Sehnen. Werden diese von Z' aus auf S' projiziert, so erhält man je zwei Elementarflächenstücke $\mathfrak{e}_1', \mathfrak{e}_2'$ und $\mathfrak{f}_1', \mathfrak{f}_2'$, die in die Sehnen von $\mathfrak{K}'$ bzw. $\mathfrak{L}'$ eingespannt sind. Die auf S' liegenden Randbogen dieser Elementarflächenstücke geben das Kreuzungsschema der Normalform k' wieder. Nach Voraussetzung gibt es eine s-Abbildung φ der S^3 auf sich, die k' in k überführt. Die Bilder von $S', \mathfrak{e}_1', \mathfrak{e}_2',$

f_1', f_2', $\mathfrak{K}'$, $\mathfrak{L}'$ bezeichnen wir durch die entsprechenden ungestrichenen Buchstaben.

Zum Beweise des Satzes wird φ so abgeändert, daß die Behauptung ersichtlich wird. Hierfür benutzen wir die Schar konzentrischer Tetraeder (T), die aus T_1 durch Dehnung mit dem Zentrum Z entsteht[17]). φ wird zunächst so gewählt werden, daß die Schar (T) zulässig bezüglich S, e_1, e_2, f_1, f_2 ist und daß e_1, e_2 keine Ausnahmeecken besitzen. Danach werden die Ausnahmeecken von S, f_1 und f_2 beseitigt (§§ 6—8), wobei wir zulassen, daß neue Ausnahmeecken auf f_1 und f_2 entstehen, wenn solche von S beseitigt werden. Nachdem alle Ausnahmeecken verschwunden sind, läßt sich zeigen (§ 9), daß S, f_1 und f_2 eines der beiden Grenztetraeder nicht durchsetzen, womit sich schließlich (§ 10) die Behauptung ergibt.

Wir beginnen damit, daß wir zunächst $\mathfrak{K}$, e_1, e_2 und danach f_1, f_2 in spezielle Lage bringen.

1. Schritt. *Normierung der Gestalt von* $\mathfrak{K}$.

Durch geeignete Änderung von φ soll $\mathfrak{K}$ folgende „wurmförmige" Gestalt erhalten:

1. Die Sehnen von $\mathfrak{K}$ liegen auf zentralen Strecken von k. Je ein Endpunkt dieser beiden Sehnen ist auf dem Rande S von $\mathfrak{K}$ ein Extremum der zentrischen Distanz[18]), während das Schartetraeder durch den anderen Sehnenendpunkt von S ein ausnahmeeckenfreies Elementarflächenstück abschneidet, für welches der eine Sehnenendpunkt mittlerer Punkt, der andere Randpunkt ist. Die beiden so erhaltenen Elementarflächenstücke auf S werden als *Köpfe* bezeichnet, ihr Rand als *Grenzlinie* des Kopfes.

2. Jeder Kopf berandet zusammen mit einem Elementarflächenstück auf dem Schartetraeder durch seine Grenzlinie eine auf $\mathfrak{K}$ liegende Kugel. Auf dieser liegt eines der beiden Elementarflächenstücke e_1, e_2, und zwar so, daß es von den Schartetraedern durch die Endpunkte der zugehörigen Sehne nur in diesen Endpunkten getroffen wird und von den dazwischen liegenden Schartetraedern in nur je einem Querschnitt.

3. $\mathfrak{K}$ liegt zwischen T_1 und T_2. Der Rand S besteht außer den beiden Köpfen aus *Knien* und *Schenkeln*. Jeder Schnitt von S mit einem Schartetraeder berandet auf diesem ein ganz auf $\mathfrak{K}$ liegendes Elementarflächenstück. Es sind also keine „Einstülpungen" vorhanden.

Um zu zeigen, daß wir φ so abändern können, bemerken wir zunächst, daß es auf $\mathfrak{K}'$ ein Elementarflächenstück e' gibt, das $\mathfrak{K}'$ so in zwei Kugeln zerlegt, daß e_1' und e_2' auf e' liegen. Man kann e' beispielsweise so erhalten, daß man die inneren Wege von k' auf T_1' zu dem Rande eines Elementarflächenstückes ergänzt und diesen Rand von Z' aus auf S' projiziert. Sei q' ein Streckenzug auf e', der die beiden Sehnen von $\mathfrak{K}'$ verbindet und mit diesen nur seine Endpunkte gemein hat. Es kann angenommen werden, daß

[17]) Vgl. hierzu und für das Folgende [6] § 3.

[18]) Zwei Punkte haben gleiche zentrische Distanz, wenn sie auf demselben Schartetraeder liegen.

die Bilder der Endpunkte von q' auf zentralen Strecken von k liegen, da sich dies durch eine s-Abbildung erreichen läßt, die k auf sich abbildet. Ferner kann angenommen werden, daß diese Punkte in einer Triangulation des Bildes von $\mathfrak{e}'$ nicht Ecken sind, andernfalls ist q' geeignet abzuändern. Das Bild von q' ist ein Streckenzug q, der bei festen Enden und festem k kombinatorisch isotop so deformiert werden kann, daß er zwischen T_1 und T_2 liegt. Es kann nun für q eine simpliziale Umgebung so gewählt werden, daß diese eine Kugel $\mathfrak{K}$ der oben beschriebenen Gestalt ausmacht, wobei der Durchschnitt von $\mathfrak{K}$ und dem Bilde von $\mathfrak{e}'$ aus einem Elementarflächenstück $\mathfrak{e}''$ besteht, dessen Durchschnitte mit den Bildern von $\mathfrak{e}_1'$ und $\mathfrak{e}_2'$ ebene Elementarflächenstücke sind, die der Forderung 2. genügen. Das Urbild von $\mathfrak{K}$ ist eine Kugel $\mathfrak{K}''$. Für $\mathfrak{K}'$ und $\mathfrak{K}''$ treffen die Voraussetzungen von Hilfssatz 1 zu, womit man die gewünschte Abänderung von φ erhält.

2. Schritt. *Normierung der Gestalt von* $\mathfrak{f}_1$ *und* $\mathfrak{f}_2$.

$\mathfrak{f}_1$ und $\mathfrak{f}_2$ werden von je einem Streckenzug r_1 bzw. r_2 auf S und je einer von k herrührenden Sehne l_1 bzw. l_2 von $\mathfrak{L}$ berandet. Bei festen $\mathfrak{e}_1$ und $\mathfrak{e}_2$ kann φ so abgeändert werden, daß S und k in sich übergehen und Folgendes gilt:

1. Für r_1 und r_2 sind Extrema der zentrischen Distanz die Endpunkte und die Schnitte mit den Rändern von $\mathfrak{e}_1$ und $\mathfrak{e}_2$. Die übrigen Extrema liegen auf Knien von S, und zwar so, daß keine zwei aufeinanderfolgenden Extrema auf demselben Knie liegen und daß r_1 und r_2 je α'-mal über jeden Schenkel von S laufen. r_1 und r_2 gehen nicht durch die Ausnahmeecken von S.

2. Soweit l_1 und l_2 mit inneren oder äußeren Wegen von k zusammenfallen, stoßen $\mathfrak{f}_1$, $\mathfrak{f}_2$ von außen an T_1 bzw. von innen an T_2 an.

3. Die Schar (T) ist zulässig in bezug auf $\mathfrak{f}_1$ und $\mathfrak{f}_2$. Diejenigen Sehnenendpunkte, die auf den Grenzlinien der Köpfe von S liegen, sind Extrema der zentrischen Distanz auf $\mathfrak{f}_1$ und $\mathfrak{f}_2$, während die übrigen Extrema der zentrischen Distanz auf r_1 und r_2 nicht zugleich Extrema auf $\mathfrak{f}_1$ und $\mathfrak{f}_2$ sind. Auf dem Rande von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ liegen keine Ausnahmeecken bezüglich (T).

4. Alle Ausnahmeecken von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ haben die Ordnung 1, wobei keine zwei Ausnahmeecken dieselbe zentrische Distanz besitzen.

Es ist noch genauer zu sagen, was unter der Zulässigkeit der Schar (T) bezüglich $\mathfrak{f}_1$ und $\mathfrak{f}_2$ in den Randpunkten von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ verstanden werden soll. Die Forderung 2. fixiert den Sachverhalt für die auf T_1 und T_2 liegenden Randpunkte. Für die übrigen ist Forderung 3. maßgebend. Für einen Randpunkt P, der nicht auf T_1 oder T_2 liegt und der nicht Extremum der zentrischen Distanz auf r_1 oder r_2 ist, wird verlangt, daß in P genau ein Schnitt von $\mathfrak{f}_1$ bzw. $\mathfrak{f}_2$ mit dem Schartetraeder durch P entspringt. Das Verhalten in den Sehnenendpunkten zeigt Fig. 9; Fig. 10 zeigt das Verhalten in einem anderen Extremum der zentrischen Distanz auf r_1 oder r_2.

Forderung 1. läßt sich dadurch erfüllen, daß man φ mit einer s-Abbildung zusammensetzt, die S in sich überführt und auf k, $\mathfrak{e}_1$ und $\mathfrak{e}_2$ die Identität ist. Betrachtet man nämlich für die Grenzlinien der Schenkel von S die Urbilder

auf S', so findet man geschlossene Wege, welche die auf S' liegenden Randstücke von e_1' und e_2' trennen und die unter Vermeidung dieser Randstücke auf S' kombinatorisch isotop gleichzeitig so deformiert werden können, daß sie die auf S' liegenden Randstücke r_1' und r_2' von f_1' und f_2' je α'-mal schneiden, und zwar dann notwendig abwechselnd, wie in Fig. 5 durch die gestrichelte Linie m angedeutet ist. Für die Grenzlinien der Köpfe ist dabei der zugehörige Sehnenendpunkt als Schnittpunkt zu zählen. Wird φ entsprechend deformiert, so treffen r_1 und r_2 jede Grenzlinie je α'-mal, und anschließend können noch r_1 und r_2 auf S kombinatorisch isotop bei festen Endpunkten und festen Schnittpunkten mit den Rändern von e_1 und e_2 so deformiert werden, daß Forderung 1. erfüllt ist.

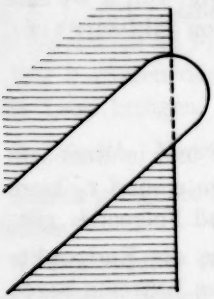


Fig. 9

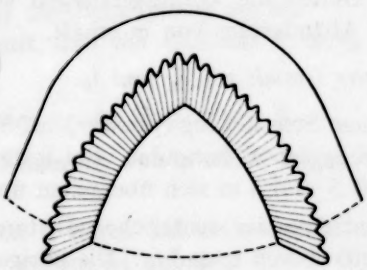


Fig. 10

Forderung 2. ist in naheliegender Weise zu erfüllen: Längs der inneren und äußeren Wege von k drille man f_1 und f_2 in die gewünschte Gestalt, wobei sich Drillungen von f_1 und f_2 um die zentralen Strecken von k ergeben. Das gewünschte Resultat läßt sich durch eine s-Abbildung erreichen, die auf $\mathfrak{R}$ die Identität ist und k in sich überführt.

Forderung 4. für die mittleren Punkte von f_1 und f_2 läßt sich durch [6] Hilfssatz 5 erfüllen, Forderung 3. für die Randpunkte zuvor durch eine entsprechende Konstruktion, wobei an Stelle des dortigen Würfels $\mathfrak{W}$ eine geeignete abgeschlossene Umgebung $\mathfrak{W}$ bezüglich $\mathfrak{L}$ für den jeweils betrachteten Punkt zu verwenden ist und der Schnitt von f_1 bzw. f_2 mit dem Rande von $\mathfrak{W}$ unter Umständen noch geeignet kombinatorisch isotop zu deformieren ist.

3. Schritt. *Veränderungen an Knien und Köpfen von* S.

Wir ändern S noch so ab, daß Folgendes gilt:

1. Haben zwei Punkte auf den Knien und Köpfen von S dieselbe zentrische Distanz, so liegen sie auf demselben Knie oder demselben Kopf.

2. Keine Ausnahmeecke von f_1 und f_2 hat dieselbe zentrische Distanz wie ein Punkt auf einem Knie oder Kopf von S.

Dabei sollen keine neuen Ausnahmeecken auf S, f_1 und f_2 erzeugt werden.

Da r_1 und r_2 nicht durch die Ausnahmeecken von S gehen, kann S in den Ausnahmeecken längs der anstoßenden Schenkel hinreichend wenig so ausgebeult werden[19]), daß keine zwei Ausnahmeecken von S dieselbe zentrische Distanz haben und daß diese zentrischen Distanzen von denjenigen der Sehnenendpunkte und der Ausnahmeecken von f_1 und f_2 verschieden sind. Außerdem kann angenommen werden, daß keine Ausnahmeecke von f_1 und f_2

[19]) Vgl. [6] § 6.

dieselbe zentrische Distanz wie ein Sehnenendpunkt hat. Die beiden Forderungen lassen sich nun dadurch erfüllen, daß die Knie und Köpfe von S bei festen Grenzlinien hinreichend stark zusammengezogen werden.

Wir beschreiben das Zusammenziehen zunächst für ein Knie. Der Anschaulichkeit halber nehmen wir an, daß das betrachtete Knie innerhalb des Projektionskegels von Z nach den „waagerechten" Tetraederseiten liegt, so daß wir die Einbettung in den R^3 benutzen können[20]). Statt von zentrischer Distanz können wir dann von Höhe sprechen. Dabei sei Z der höchste Punkt des Projektionskegels. Ein inneres Knie von S kann als „oberes Knie" aufgefaßt werden. Zur Fixierung der Sprechweise legen wir ein solches zugrunde.

Das betrachtete Knie $\mathfrak{k}$ liegt oberhalb der waagerechten Ebene E durch seine Grenzlinien. Von den auf $\mathfrak{k}$ liegenden Stücken von r_1 und r_2 betrachten wir zuerst dasjenige, es heiße t, das die geringste Höhe erreicht. t kann über ein Elementarflächenstück $\mathfrak{g}$ auf $\mathfrak{k}$ so deformiert werden, daß t beliebig nahe an E zu liegen kommt. Dabei wird $\mathfrak{g}$ so gewählt, daß es von jeder zu E parallelen Ebene in höchstens einem Querschnitt getroffen wird, was offenbar möglich ist. Gehört t zum Rande von $\mathfrak{f}_1$, so läßt sich $\mathfrak{f}_1$ durch $\mathfrak{g}$ ergänzen und anschließend $\mathfrak{g}$ von $\mathfrak{k}$ bei festem t abheben, indem man die Ecken einer geeigneten Triangulation von $\mathfrak{f}_1 + \mathfrak{g}$, die auf $\mathfrak{g}$, nicht aber auf t liegen, hinreichend wenig waagerecht ins Äußere von $\mathfrak{R}$ verschiebt. Das Resultat läßt sich auch durch eine s-Abbildung erreichen, die S in sich überführt und außerhalb einer beliebig kleinen Umgebung von $\mathfrak{g}$ die Identität ist. Es können so der Reihe nach alle auf $\mathfrak{k}$ liegenden Stücke von r_1 und r_2 hinreichend nahe an E herandeformiert werden, wobei nur das jeweils benutzte Elementarflächenstück oberhalb der bereits deformierten Stücke von r_1 und r_2 liegen muß. Nachdem dies geschehen ist, kann $\mathfrak{k}$ bei festem $\mathfrak{f}_1$ und $\mathfrak{f}_2$ so deformiert werden, daß $\mathfrak{k}$ ein Knie bleibt und ganz unterhalb einer waagerechten Ebene liegt, die beliebig dicht oberhalb des höchsten Punktes der auf $\mathfrak{k}$ liegenden Stücke von r_1 und r_2 verläuft.

Für einen Kopf $\mathfrak{k}$ von S benutzen wir entsprechend die Einbettung in den R^3, und wir nehmen an, daß $\mathfrak{k}$ oberhalb der waagerechten Ebene E durch seine Grenzlinie liegt. Innerhalb $\mathfrak{k}$ liegt eines der beiden Elementarflächenstücke $\mathfrak{e}_1$, $\mathfrak{e}_2$, etwa $\mathfrak{e}_1$. Ähnlich wie im Falle eines Knies können hier die auf $\mathfrak{k}$ liegenden Stücke von r_1 und r_2 an E herandeformiert werden bis auf dasjenige, das im höchsten Punkte von $\mathfrak{k}$ endet. Nur ist hier die entsprechende Abbildung so vorzunehmen, daß $\mathfrak{e}_1$ in sich übergeht und die Schnitte von r_1 und r_2 mit dem Rande von $\mathfrak{e}_1$ Maxima der Höhe auf r_1 und r_2 bleiben, was bei geeigneter Wahl der zur Deformation benutzten Elementarflächenstücke möglich ist. Das Zusammenziehen des Kopfes kann danach so geschehen, daß k auf sich und $\mathfrak{e}_1$ in sich abgebildet wird und daß das im höchsten Punkte von $\mathfrak{k}$ endende Stück von r_1 oder r_2 monoton bleibt.

[20]) Vgl. den wiederholt benutzten Sachverhalt in [6].

§ 6. Beseitigung von Ausnahmeecken auf $\mathfrak{f}_1$ und $\mathfrak{f}_2$

Wir betrachten alle Schartetraeder durch die Ausnahmeecken von $\mathfrak{f}_1$ und für jedes solche Schartetraeder die Schnitte mit $\mathfrak{f}_1$, die durch die betreffende Ausnahmeecke gehen. Diese Schnitte sind geschlossen oder gehen zum Rande von $\mathfrak{f}_1$. Unter den geschlossenen gibt es einen auf $\mathfrak{f}_1$ innersten, der auf $\mathfrak{f}_1$ ein von weiteren Schnitten freies Elementarflächenstück, also eine *Haube*, berandet. Wir werden zunächst alle Hauben beseitigen.

Die danach verbleibenden Schnitte durch Ausnahmeecken von $\mathfrak{f}_1$ mit Schartetraedern führen sämtlich zum Rande von $\mathfrak{f}_1$. Dabei gehen von jeder Ausnahmeecke, wegen der Ordnung 1, vier solche Schnitte aus, wodurch $\mathfrak{f}_1$ jeweils in vier Elementarflächenstücke zerlegt wird. Wir nennen ein solches Elementarflächenstück einen *Zipfel*, wenn auf ihm keine weitere Ausnahmeecke liegt.

Ein Zipfel von $\mathfrak{f}_1$ hat mit dem Rande von $\mathfrak{f}_1$ ein Stück t gemein, das entweder auf S oder auf k oder teilweise auf S und teilweise auf k liegt. Es verschwinden alle Ausnahmeecken, wenn man der Reihe nach diejenigen beseitigt, von denen ein zu k punktfremder Zipfel ausgeht. Andernfalls gäbe es dann zwei von derselben Ausnahmeecke ausgehende Zipfel, die in einem Schnitt aneinanderstoßen und S nicht treffen. Das auf dem Rande der beiden Zipfel liegende Stück von k ließe sich über die beiden Zipfel deformieren, und man erhielte, daß k in eine Brückendarstellung mit nur einer Brücke verwandelt werden könnte, also den Kreis darstellte, was ausgeschlossen wurde.

Wir brauchen also nur solche Zipfel zu betrachten, bei denen das mit t bezeichnete Randstück ganz auf S liegt. Wegen des 3. Schrittes von § 5 kann dabei t nicht ganz auf einem Knie oder Kopf von S liegen. Wir betrachten hier nur solche Zipfel, bei denen das Randstück t die Köpfe von S nicht trifft, und behandeln die verbleibenden in § 8.

1. Schritt. *Beseitigung von Hauben.*

Da wir ausgeschlossen haben, daß k zwei unverkettete Kreise darstellt, ist eine Kugel, deren Rand k nicht trifft, entweder punktfremd zu k oder sie enthält k im Inneren. Hierauf beruht es, daß wir die Hauben von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ nach dem Verfahren von [6] § 5 beseitigen können. Zunächst werden wie dort im 1. Schritt alle diejenigen Ausnahmeecken von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ beseitigt, durch welche eine geschlossene Schnittlinie von $\mathfrak{f}_1$ bzw. $\mathfrak{f}_2$ mit dem zugehörigen Schartetraeder geht, die auf dem Schartetraeder ein zu k und S punktfremdes Elementarflächenstück berandet. Insbesondere verschwinden dabei alle Ausnahmeecken von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ innerhalb T_1 und außerhalb T_2, und es bleiben k und S ungeändert.

Die weitere Beseitigung von Hauben verläuft entsprechend dem 2. bis 5. Schritt von [6] § 5. Bei dem Ausschieben (5. Schritt) ist zu beachten, daß der dort mit h bezeichnete Streckenzug k und S nicht und $\mathfrak{f}_1$ und $\mathfrak{f}_2$ nur in solchen Punkten treffen darf, die Minima bzw. Maxima der zentrischen Distanz auf $\mathfrak{f}_1$ und $\mathfrak{f}_2$ sind, je nachdem eine innere oder äußere Haube

beseitigt wird. Diese Wahl von h ist möglich wegen Forderung 3. des 2. Schrittes in § 5.

Bei diesem Vorgang ändert sich die Gestalt von k. Die ursprüngliche Gestalt von k läßt sich jedoch durch eine zentrische Abbildung wieder herstellen: Zur Fixierung der Sprechweise nehmen wir an, daß eine innere Haube beseitigt wurde. Man kann dann zunächst alle bei der Beseitigung der Haube benutzten zentrischen Hilfsabbildungen rückgängig machen, so daß also die Haube durch eine s-Abbildung χ beseitigt wird, die außerhalb T_2 die Identität ist. Nach dem Verfahren von [6] Hilfssatz 1 kann χ durch eine zentrische Abbildung, die im Komplement von T_2 die Identität ist, so abgeändert werden, daß die zentralen Strecken von k in zentrale Strecken übergehen, also festbleiben. k und $\chi(k)$ sind dann Brückendarstellungen mit zwei Brücken im Sinne von § 1 und stimmen in den zentralen Strecken und den äußeren Wegen überein. Die Brücken von k und $\chi(k)$ bilden in der von T_2 berandeten Kugel, die das Innere von T_2 ausmacht, je zwei einfache Sehnen. In die von k herrührenden Sehnen spannen wir Elementarflächenstücke $\mathfrak{m}_1$, $\mathfrak{m}_2$ ein, indem wir die Brücken von k von Z aus auf T_2 projizieren. Entsprechend werden die Elementarflächenstücke $\mathfrak{n}_1$, $\mathfrak{n}_2$ in die von $\chi(k)$ herrührenden Sehnen eingespannt. In diese Sehnen können aber auch die Elementarflächenstücke $\chi(\mathfrak{m}_1)$, $\chi(\mathfrak{m}_2)$ eingespannt werden. Nach Hilfssatz 3 sind die auf T_2 liegenden Randstücke von $\mathfrak{n}_1$, $\mathfrak{n}_2$ zu den entsprechenden Randstücken von $\chi(\mathfrak{m}_1)$, $\chi(\mathfrak{m}_2)$ kombinatorisch isotop auf T_2. Da die bezeichneten Randstücke Zentralprojektionen der inneren Wege von $\chi(k)$ und k sind, lassen sich die inneren Wege von $\chi(k)$ auf T_1 kombinatorisch isotop in die inneren Wege von k deformieren, wodurch $\chi(k)$ in k übergeht. Das Deformationsresultat läßt sich auch durch eine zentrische Abbildung erhalten[21]).

2. Schritt. *Beseitigung von Zipfeln, die k und die Köpfe von S nicht treffen.*

Es sei $\mathfrak{z}$ ein Zipfel der genannten Art, etwa auf $\mathfrak{f}_1$. $\mathfrak{z}$ wird berandet von zwei Schnitten s_1, s_2 von $\mathfrak{f}_1$ mit einem Schartetraeder T durch eine Ausnahmeecke A von $\mathfrak{f}_1$, wobei s_1 und s_2 von A ausgehen und zum Rande von $\mathfrak{f}_1$ führen, und einem Randstück t von $\mathfrak{f}_1$, das auf S liegt und die Köpfe von S nicht trifft. Nach der Gestalt von t unterscheiden wir zwei Fälle.

Fall a. t liegt auf einem Knie und den beiden anstoßenden Schenkeln von S.

Das Verfahren besteht darin, daß wir das Stück von S, auf dem t verläuft, vermöge $\mathfrak{z}$ so abbiegen, daß A Extremum der zentrischen Distanz des Randes von $\mathfrak{f}_1$ wird und auf ein Knie von S zu liegen kommt. Es entstehen dabei neue Ausnahmeecken, die jedoch nach dem Verfahren des 1. Schrittes beseitigt werden können.

t liege auf dem Knie $\mathfrak{k}$ und den beiden anstoßenden Schenkeln $\mathfrak{s}_1$ und $\mathfrak{s}_2$ von S. Zur Fixierung der Sprechweise nehmen wir an, daß $\mathfrak{k}$ ein inneres Knie ist. Durch eine zentrische Abbildung kann erreicht werden, daß $\mathfrak{k}$, $\mathfrak{s}_1$, $\mathfrak{s}_2$ und $\mathfrak{z}$ im Inneren des Projektionskegels von Z nach den „waagerechten"

[21]) Vgl. die Konstruktion in [7] S. 144.

Tetraederseiten liegen, so daß wir die Einbettung in den R^3 wie im 3. Schritt von § 5 benutzen und von Höhe statt von zentrischer Distanz sprechen können. Der Sattelpunkt A von $\mathfrak{f}_1$ liegt nun in einer waagerechten Ebene E, in der s_1 und s_2 verlaufen, und es liegt $\mathfrak{z}$ oberhalb E. Der höchste Punkt H von t ist nach dem 2. Schritt von § 5 nicht zugleich höchster Punkt von $\mathfrak{z}$. Daher wird $\mathfrak{z}$ von der waagerechten Ebene durch H in einer geschlossenen Schnittlinie geschnitten und besitzt oberhalb dieser Ebene haubenförmige Gestalt. Durch eine s-Abbildung, die außerhalb einer beliebig kleinen Umgebung von $\mathfrak{z}$, insbesondere auf k, $\mathfrak{f}_2$, e_1, e_2 und den Köpfen von S die Identität ist und die $\mathfrak{f}_1$ in sich abbildet, kann Folgendes erreicht werden:

Die von S berandete Kugel $\mathfrak{K}$ wird längs $\mathfrak{z}$ und dann noch etwas weiter unterhalb s_1 und s_2 ausgebeult, wobei gleichzeitig $\mathfrak{f}_1$ in sich zusammengezogen wird, und zwar so, daß der ausgebeulte Teil von $\mathfrak{K}$ hinreichend dicht unter $\mathfrak{z}$ liegt und daß auf S nur zwei neue Ausnahmeecken entstehen, nämlich der höchste Punkt H von t und eine neue Ausnahmeecke Q dicht unterhalb A. Das auf $\mathfrak{s}_1$, $\mathfrak{k}$ und $\mathfrak{s}_2$ liegende Stück von r_1, das t umfaßt, geht dabei in ein solches durch A über, wobei A das einzige Maximum der Höhe für dieses Stück ist.

Ähnlich wie im 3. Schritt von § 5 werden nun die übrigen auf $\mathfrak{s}_1$, $\mathfrak{k}$, $\mathfrak{s}_2$ verlaufenden Stücke von r_1 und r_2 mittels geeigneter Elementarflächenstücke auf S so deformiert, daß auf ihnen die Höhe wieder nur ein einziges Maximum annimmt und daß dieses auf dem ausgebeulten Stück hinreichend dicht über A liegt. Bei dieser Deformation werden die anstoßenden Stücke von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ „mitgezogen". Es entstehen dabei neue Ausnahmeecken auf $\mathfrak{f}_1$ und $\mathfrak{f}_2$. Die zur Deformation benutzten Elementarflächenstücke können jedoch so gewählt werden, daß sie die Ausnahmeecke in den Grenzlinien des Knies $\mathfrak{k}$ vermeiden. Kann ein solches Elementarflächenstück zudem so gewählt werden, daß es auch den höchsten Punkt von $\mathfrak{k}$ vermeidet, so läßt es sich so bestimmen, daß auf dem mitgezogenen Stück von $\mathfrak{f}_1$ bzw. $\mathfrak{f}_2$ nur eine einzige Ausnahmeecke entsteht, die dicht bei H liegt. Muß das benutzte Elementarflächenstück so gewählt werden, daß der höchste Punkt von $\mathfrak{k}$ auf ihm liegt, so läßt es sich einrichten, daß auf dem entsprechend mitgezogenen Stück von $\mathfrak{f}_1$ bzw. $\mathfrak{f}_2$ nur zwei Ausnahmeecken entstehen, die eine dicht bei H, die andere dicht bei dem bisherigen Maximum der Höhe des Randbogens. Durch die neu entstehenden Ausnahmeecken von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ gehen geschlossene Schnittlinien von $\mathfrak{f}_1$ bzw. $\mathfrak{f}_2$ mit den zugehörigen Schartetraedern, wie sich aus der Gestalt von $\mathfrak{k}$ und der Gestalt von $\mathfrak{z}$ oberhalb der waagerechten Ebene durch H ergibt, und können anschließend wieder nach dem Verfahren des 1. Schrittes beseitigt werden. Zuvor ist noch das oberhalb E liegende, ausgebeulte Stück von $\mathfrak{K}$ dicht an E heranzuziehen, so daß $\mathfrak{f}_1$ und $\mathfrak{f}_2$ festbleiben und Q als einzige Ausnahmeecke des verbeulten Stückes verbleibt. Dies ist möglich, da eine waagerechte Ebene hinreichend dicht über E das ausgebeulte Stück von $\mathfrak{K}$ in nur einem Elementarflächenstück schneidet.

Das geschilderte Verfahren soll für $\mathfrak{f}_1$ und $\mathfrak{f}_2$ so oft wie möglich angewandt werden.

Fall b. t verläuft über mindestens zwei Knie von S.

Das Schartetraeder T durch die Ausnahmeecke A schneidet S in zwei Schnitten u_1 und u_2, auf denen s_1 und s_2 enden. u_1 und u_2 schneiden aus S einen Kreisring $\mathfrak{s}$ aus, der t enthält und auf dem mindestens zwei Knie liegen. Auf T beranden u_1 und u_2 zwei auf $\mathfrak{K}$ liegende Elementarflächenstücke, die zusammen mit $\mathfrak{s}$ eine auf $\mathfrak{K}$ liegende Kugel $\mathfrak{M}$ beranden. Durch eine s-Abbildung, die $k, \mathfrak{e}_1, \mathfrak{e}_2$ und die Köpfe von S festläßt, kann $\mathfrak{M}$ längs des Zipfels $\mathfrak{z}$ so abgebogen werden, daß s_1 und s_2 ins Innere von $\mathfrak{K}$ geraten, daß S weiterhin aus Knien, Schenkeln und den beiden Köpfen besteht und daß an Stelle der auf $\mathfrak{s}$ liegenden Knie kein oder ein Knie tritt, je nachdem die Anzahl der Knie auf $\mathfrak{s}$ gerade oder ungerade ist, während die übrigen Knie unverändert erhalten bleiben. Die Zahl der Ausnahmeecken von S vermindert sich auf diese Weise um mindestens zwei. Das Beweisverfahren ist hierauf vom 2. Schritt des § 5 an zu wiederholen, und zwar so oft, bis Zipfel der hier betrachteten Art auf $\mathfrak{f}_1$ und $\mathfrak{f}_2$ nicht mehr vorhanden sind.

§ 7. Beseitigung der Ausnahmeecken von S

Wir haben soeben einen Sachverhalt betrachtet, der zu einer Reduktion der Ausnahmeecken von S Anlaß gab. Ähnliche Reduktionen können möglich sein, wenn man an Stelle von Ausnahmeecken die Extrema der zentrischen Distanz auf r_1 und r_2 betrachtet. Das Schartetraeder durch ein solches Extremum A, das nicht Sehnenendpunkt ist, schneidet $\mathfrak{f}_1$ bzw. $\mathfrak{f}_2$ entweder in einer geschlossenen Schnittlinie durch A oder in zwei von A ausgehenden Querschnitten. Im ersten Falle nennen wir das *Extremum A regulär*, im zweiten *singulär*. Zwei der Sehnenendpunkte sind Extrema der zentrischen Distanz auf $\mathfrak{f}_1$ oder $\mathfrak{f}_2$. Wir bezeichnen auch diese beiden Punkte als reguläre Extrema der zentrischen Distanz auf r_1 und r_2. Die beiden anderen Sehnenendpunkte zählen zu den singulären Extrema. Das Schartetraeder durch einen solchen Punkt schneidet $\mathfrak{f}_1$ oder $\mathfrak{f}_2$ in einem Querschnitt, der von diesem Punkte ausgeht. Wir beseitigen zunächst Ausnahmeecken von S vermöge der Querschnitte auf $\mathfrak{f}_1$ und $\mathfrak{f}_2$, die von singulären Extrema ausgehen, und danach die restlichen Ausnahmeecken von S. Nach jeder solchen Reduktion ist das Beweisverfahren erneut mit dem 2. Schritt von § 5 zu beginnen.

1. Schritt. *Reduktionen vermöge singulärer Extrema auf Knien.*

Sei A ein solches Extremum der zentrischen Distanz, etwa auf r_1, und es liege A auf dem Knie $\mathfrak{k}$ von S. Das Schartetraeder T durch A schneidet $\mathfrak{f}_1$ in zwei Querschnitten s_1 und s_2, die von A ausgehen. s_1 und s_2 schneiden von $\mathfrak{f}_1$ zwei Elementarflächenstücke $\mathfrak{g}_1$ und $\mathfrak{g}_2$ ab, die von s_1 bzw. s_2 und einem Stück t_1 bzw. t_2 des Randes von $\mathfrak{f}_1$ berandet werden. Eine Reduktion ist möglich, wenn eines der beiden Randstücke t_1, t_2 entweder keinen der beiden Köpfe von S trifft oder nur einen von beiden und dann ein Stück mit k gemein hat. Es liege etwa für t_1 einer dieser beiden Fälle vor.

Fall a. t_1 *trifft keinen der beiden Köpfe von* S.

In diesem Falle liegt t_1 auf S. Seien v_1 und v_2 die Grenzlinien des Knies $\mathfrak{k}$. Nur eine von beiden, etwa v_2, wird von t_1 getroffen, und zwar nur einmal. Wegen des 3. Schrittes von § 5 endet t_1 auf einem Schnitte u von T mit einem Schenkel von S im Endpunkt von s_1. Aus S wird von v_1 und u ein Kreisring $\mathfrak{s}$ ausgeschnitten, auf dem t_1 verläuft. Auf $\mathfrak{s}$ liegt mindestens ein von $\mathfrak{k}$ verschiedenes Knie, und es berandet $\mathfrak{s}$ zusammen mit je einem Elementarflächenstück auf den Schartetraedern durch v_1 und u eine auf $\mathfrak{R}$ liegende Kugel $\mathfrak{M}$. Wie in § 6, 2. Schritt, Fall b kann nun $\mathfrak{M}$ längs $\mathfrak{g}_1$ so abgebogen werden, daß s_1 ins Innere von $\mathfrak{M}$ gelangt, S weiterhin aus Knien, Schenkeln und den beiden Köpfen besteht und mindestens zwei Knie von S verschwinden.

Fall b. t_1 *hat ein Stück mit* k *gemein und trifft nur einen der beiden Köpfe von* S.

Sind v_1 und v_2 wieder die Grenzlinien von $\mathfrak{k}$, so wird auch in diesem Falle nur eine von beiden, etwa v_2, von t_1 getroffen. v_1 zerlegt S in zwei Elementarflächenstücke. Auf einem von beiden, es heiße $\mathfrak{s}$, liegt ein Stück von t_1, während der Rest von t_1 ein Stück von k ausmacht. v_1 berandet auf seinem Schartetraeder ein Elementarflächenstück, das zusammen mit $\mathfrak{s}$ eine auf $\mathfrak{R}$ liegende Kugel $\mathfrak{M}$ berandet. Entsprechend dem Vorangehenden wird $\mathfrak{M}$ längs $\mathfrak{g}_1$ abgebogen, so daß $\mathfrak{s}$ in einen Kopf und ein Stück eines Schenkels übergeht. Der auf $\mathfrak{s}$ liegende Kopf ist hierbei längs k zu verschieben, möglicherweise über innere und äußere Wege von k hinweg. Die hierzu benötigte s-Abbildung ist so zu wählen, daß k in sich übergeht und daß das an den Kopf von $\mathfrak{s}$ anstoßende Elementarflächenstück $\mathfrak{e}_1$ oder $\mathfrak{e}_2$ wieder die geforderte Gestalt (§ 5, 1. Schritt) erhält.

2. Schritt. *Reduktionen vermöge singulärer Extrema auf den Köpfen.*

Sei A ein solches Extremum der zentrischen Distanz, etwa auf r_1. Ist A nicht Sehnenendpunkt, so gehen von A zwei Querschnitte s_1 und s_2 aus, die auf dem Schartetraeder T durch A liegen. Eine Reduktion ist dann möglich, wenn einer dieser beiden Querschnitte von $\mathfrak{f}_1$ ein Elementarflächenstück abschneidet, das k und den zweiten Kopf von S nicht trifft. Sei dies etwa für s_1 der Fall. s_1 schneidet von $\mathfrak{f}_1$ ein Elementarflächenstück $\mathfrak{g}_1$ ab, das von s_1 und einem Randstück t_1 von $\mathfrak{f}_1$ berandet wird. t_1 trifft k und den zweiten Kopf von S nicht und endet auf einem Schnitt u von T mit einem Schenkel von S. Auf S berandet u ein Elementarflächenstück $\mathfrak{s}$, das t_1 enthält und auf dem mindestens ein Knie liegt. $\mathfrak{s}$ berandet zusammen mit einem Elementarflächenstück auf T eine auf $\mathfrak{R}$ liegende Kugel $\mathfrak{M}$. Wie oben wird $\mathfrak{M}$ längs $\mathfrak{g}_1$ so abgebogen, daß $\mathfrak{s}$ in einen Kopf und ein Stück eines Schenkels übergeht. Dies kann so geschehen, daß k in sich und das an den Kopf von $\mathfrak{s}$ anstoßende Elementarflächenstück $\mathfrak{e}_1$ oder $\mathfrak{e}_2$ wieder in ein solches mit den geforderten Eigenschaften übergeht.

Ist A Sehnenendpunkt, so wird $\mathfrak{f}_1$ von dem Schartetraeder durch A in einem von A ausgehenden Querschnitt s getroffen. Wenn s von $\mathfrak{f}_1$ ein Elementarflächenstück abschneidet, das den zweiten Kopf von S nicht trifft

und das mit k nur den Punkt A gemein hat, so kann eine Reduktion ausgeführt werden, die der eben beschriebenen entspricht.

3. Schritt. *Vorbereitung für die weitere Reduktion.*

Es sei keine der vorangehenden Reduktionen mehr ausführbar aber noch mindestens ein Knie von S vorhanden. Wir bestimmen ein Elementarflächenstück $\mathfrak{g}$, vermöge dessen eine weitere Reduktion möglich ist.

Nehmen wir zunächst an, daß auf $\mathfrak{f}_1$ noch Ausnahmeecken vorhanden sind. Für mindestens eine Ausnahmeecke A tritt dann der Fall ein, daß von A zwei „benachbarte" Zipfel ausgehen, die einen von A ausgehenden Schnitt von $\mathfrak{f}_1$ mit dem Schartetraeder durch A als gemeinsames Randstück besitzen und von denen mindestens einer zu k punktfremd ist. Wenn möglich, wählen wir A so, daß beide Zipfel zu k punktfremd sind, und betrachten den dann vorliegenden Sachverhalt.

Fall a. Auf $\mathfrak{f}_1$ existieren zwei benachbarte Zipfel $\mathfrak{z}_1$ und $\mathfrak{z}_2$, die k nicht treffen.

$\mathfrak{z}_1$ und $\mathfrak{z}_2$ haben mit r_1 je ein Stück des Randes t_1 bzw. t_2 gemein. Wegen § 6 führen t_1 und t_2 mindestens je einmal über einen Kopf. Da t_1 und t_2 einen Endpunkt gemein haben und mindestens ein Knie von S vorhanden ist, gibt es auf t_1 oder t_2 mindestens ein Extremum der zentrischen Distanz, das auf einem Knie liegt. Sei dies etwa für t_1 der Fall. Wir betrachten alle Schnitte von $\mathfrak{z}_1$ mit Schartetraedern, die durch die Extrema der zentrischen Distanz von t_1 gehen. Sind unter diesen Querschnitte, so gibt es einen solchen q, der von $\mathfrak{z}_1$ ein Elementarflächenstück $\mathfrak{g}$ abschneidet, das keinen weiteren Querschnitt enthält und das von q und einem Stück t_1' von t_1 berandet wird. Sollte kein Querschnitt vorhanden sein, was sich als unmöglich erweisen wird, so setzen wir $\mathfrak{g}=\mathfrak{z}_1$ und $t_1'=t_1$. Wegen der vorangehenden Reduktionen geht t_1' mindestens einmal über einen Kopf $\mathfrak{k}_1$, und es sei A_1 ein Extremum der zentrischen Distanz von t_1', das auf $\mathfrak{k}_1$ liegt. Nach Wahl von q ist A_1 regulär. Das Schartetraeder T_1' durch A_1 schneidet $\mathfrak{g}$ in einer geschlossenen Schnittlinie s_1 durch A_1. $\mathfrak{s}$ sei der an $\mathfrak{k}_1$ anstoßende Schenkel, $\mathfrak{k}_2$ das folgende Knie. Wir dehnen T_1' stetig in dem Sinne[22]), daß aus dem Schnitt von T_1' mit $\mathfrak{k}_1$ ein solcher mit $\mathfrak{s}$ und schließlich ein solcher mit $\mathfrak{k}_2$ entsteht, und verfolgen dabei den aus s_1 entstehenden Schnitt mit $\mathfrak{g}$. Dieser geht zunächst in einen Querschnitt von $\mathfrak{g}$ über, der auf $\mathfrak{k}_1$ entspringt und endet, danach in einen solchen, der auf $\mathfrak{s}$ entspringt und endet, schließlich in einen solchen, dessen Endpunkte auf $\mathfrak{k}_2$ liegen, bis wir zu einem Extremum der zentrischen Distanz von t_1' auf $\mathfrak{k}_2$ gelangen. Dieser Sachverhalt muß nach Wahl von $\mathfrak{g}$ eintreten, weil auf $\mathfrak{g}$ außer q keine Querschnitte durch Extrema der zentrischen Distanz von t_1 und keine Querschnitte durch Ausnahmeecken liegen. Bei stetigem Dehnen von T_1' im angegebenen Sinne geht also s_1 in einen Querschnitt von $\mathfrak{z}_1$ über, der von einem Extremum A_2 der zentrischen Distanz von t_1 auf $\mathfrak{k}_2$ ausgeht, und dieser Querschnitt muß q sein. Nach Wahl von t_1 bemerkt man zugleich, daß die Annahme $\mathfrak{g}=\mathfrak{z}_1$ widerspruchsvoll ist. Das Auffinden des

[22]) Es sind Dehnungsfaktoren kleiner als 1 zugelassen.

Elementarflächenstückes $\mathfrak{g}$, dessen Gestalt sich aus dem Vorangehenden ergibt, war unser Ziel.

Fall b. Sind $\mathfrak{z}_1$ und $\mathfrak{z}_2$ zwei benachbarte Zipfel auf $\mathfrak{f}_1$, so trifft einer von beiden k.

In diesem Falle können $\mathfrak{z}_1$ und $\mathfrak{z}_2$ so gewählt werden, daß $\mathfrak{z}_1$ zu k punktfremd ist, während auf dem Rande von $\mathfrak{z}_2$ ein Sehnenendpunkt P liegt. $\mathfrak{z}_1$ bzw. $\mathfrak{z}_2$ hat mit dem Rande von $\mathfrak{f}_1$ ein Stück t_1 bzw. t_2 gemein. t_1 geht über einen Kopf von S; geht t_1 auch über ein Knie, so kann auf $\mathfrak{z}_1$ ein Elementarflächenstück $\mathfrak{g}$ wie im Falle a gefunden werden. Andernfalls geht t_2 über alle Knie. Wir betrachten dann alle Schnitte von $\mathfrak{z}_2$ mit Schartetraedern durch Extrema der zentrischen Distanz von t_2. Gibt es darunter einen Querschnitt, der von $\mathfrak{z}_2$ ein zu k punktfremdes Elementarflächenstück abschneidet, so wird das Elementarflächenstück $\mathfrak{g}$ wie im Falle a bestimmt. Die Annahme, daß es keinen solchen Querschnitt von $\mathfrak{z}_2$ gibt, führt zu einem Widerspruch: Wir betrachten das Schartetraeder durch P und dehnen es stetig in dem Sinne, daß der entstehende Querschnitt auf $\mathfrak{z}_2$ von $\mathfrak{z}_2$ ein Elementarflächenstück abschneidet, das P als Randpunkt besitzt, nicht aber die Ausnahmeecke auf dem Rande von $\mathfrak{z}_2$. Nach Annahme könnte dies solange geschehen, bis man erstmals zu einem von P verschiedenen Extremum der zentrischen Distanz von t_2 gelangte, womit sich ein Widerspruch zu Fall b des 1. Schrittes ergibt.

Fall c. Auf $\mathfrak{f}_1$ liegen keine Ausnahmeecken.

Das Elementarflächenstück $\mathfrak{g}$ kann wie im Falle b gefunden werden, wobei $\mathfrak{f}_1$ an die Stelle von $\mathfrak{z}_2$ tritt.

4. Schritt. *Beseitigung der restlichen Ausnahmeecken von S.*

$\mathfrak{g}$, $\mathfrak{k}_1$, $\mathfrak{s}$, $\mathfrak{k}_2$ mögen dieselbe Bedeutung wie im 3. Schritt haben, ebenso t_1' und q. Das Schartetraeder durch q sei T_2'. Es schneidet $\mathfrak{k}_2$ in einer geschlossenen Schnittlinie. Ein Stück q' dieses Schnittes berandet zusammen mit t_1' ein Elementarflächenstück $\mathfrak{m}$, das auf dem von $\mathfrak{k}_1$, $\mathfrak{s}$ und $\mathfrak{k}_2$ gebildeten „Endstück" von S liegt und das T_2' nicht durchsetzt. Die Ausnahmeecke von $\mathfrak{k}_2$ liegt daher nicht auf $\mathfrak{m}$. Es beranden $\mathfrak{g}$, $\mathfrak{m}$ und ein von q und q' auf T_2' berandetes Elementarflächenstück $\mathfrak{n}$ zusammen eine Kugel $\mathfrak{M}$, deren Durchschnitt mit dem „Endstück" von S aus $\mathfrak{m}$ besteht und die mit $\mathfrak{n}$ an T_2' ohne Durchdringung anstößt. Vermöge $\mathfrak{M}$ werden wir das Endstück von S so abbiegen, daß das Knie $\mathfrak{k}_2$ verschwindet.

Zur Fixierung der Sprechweise nehmen wir an, daß $\mathfrak{k}_2$ ein äußeres Knie ist. $\mathfrak{M}$ liegt dann innerhalb T_2'. Auf $\mathfrak{m}$ liegt ein Schnittpunkt Q von S und k, der Endpunkt eines auf $\mathfrak{M}$ liegenden Stückes k_1 von k ist. k_1 liegt entweder auf einer zentralen Strecke von k, oder es besteht k_1 aus einem inneren Wege von k und je einem Stück der anstoßenden zentralen Strecken.

Fall a. k_1 liegt auf einer zentralen Strecke von k.

In diesem Falle liegt k_1 zwischen T_2' und dem Schartetraeder durch Q. Wir betrachten die auf $\mathfrak{M}$ liegenden Stücke von S und lassen $\mathfrak{f}_1$ und $\mathfrak{f}_2$ außer Betracht. Durch Deformation von S erreichen wir zunächst, daß kein äußeres

9272-65

Knie von S auf $\mathfrak{M}$ liegt: Man kann diese Knie zunächst längs geeigneter, zentrisch monotoner Streckenzüge so ausbeulen[23]), daß die Maxima der zentrischen Distanz außerhalb T_2' gelangen, und anschließend von ihrer Ausnahmeecke her (in das Knie hinein) so, daß die Ausnahmeecken ebenfalls außerhalb T_2' gelangen[24]).

Wenn nun auf k_1 außer dem Endpunkt Q kein weiterer Schnittpunkt mit S liegt, so läßt sich ein Elementarflächenstück $\mathfrak{a}$ so bestimmen[25]), daß erstens $\mathfrak{a}$ berandet wird von k_1, einem zentrisch monotonen Streckenzug m auf $\mathfrak{m}$ von Q nach einem Punkte von T_2' und von einem Streckenzug auf T_2', der die Endpunkte von m und k_1 verbindet, daß zweitens S und k von $\mathfrak{a}$ nur in m und k_1 getroffen werden und daß drittens $\mathfrak{a}$ von jedem Schartetraeder zwischen T_2' und Q in nur einem Querschnitt und von dem Schartetraeder durch Q nur in Q getroffen wird. Daß $\mathfrak{a}$ so bestimmt werden kann, erkennt man, wenn man das Schartetraeder durch Q stetig dehnt, weil die Ausnahmeecke von $\mathfrak{k}_2$ nicht auf $\mathfrak{M}$ liegt. Das Knie $\mathfrak{k}_2$ läßt sich nun dadurch beseitigen, daß man das Endstück von S längs $\mathfrak{a}$ entsprechend dem Fall b des 1. Schrittes abbiegt.

Wenn auf k_1 außer Q noch Schnittpunkte mit S liegen, so sind es die Schnitte von k mit dem von $\mathfrak{k}_1$ verschiedenen Kopf $\mathfrak{k}'$ von S. Wenn der an $\mathfrak{k}'$ anstoßende Schenkel nach außen führt, so schneidet er T_2', weil auf $\mathfrak{M}$ keine äußeren Knie mehr liegen, und es kann $\mathfrak{k}'$ nach dem eben beschriebenen Verfahren längs k_1 außerhalb $\mathfrak{M}$ geschoben werden, worauf das aus $\mathfrak{k}_1$, $\mathfrak{s}_1$ und $\mathfrak{k}_2$ bestehende Endstück von S wie oben abgebogen werden kann. Wenn der an $\mathfrak{k}'$ anstoßende Schenkel nach innen führt, so läßt sich $\mathfrak{k}'$ längs k_1 ebenfalls aus $\mathfrak{M}$ herausschieben, worauf $\mathfrak{a}$ wie oben bestimmt werden kann.

Fall b. k_1 umfaßt einen inneren Weg von k, hat aber mit S nur einen Endpunkt gemein.

Der Endpunkt Q von k_1 ist Minimum der zentrischen Distanz auf dem Kopf $\mathfrak{k}_1$. Wir beulen zunächst $\mathfrak{k}_1$ längs k_1 so aus, daß der Sehnenendpunkt Q auf die zweite zentrale Strecke von k_1 gerät, daß auf dem ausgebeulten Stück nur ein Extremum der zentrischen Distanz, und zwar ein Minimum innerhalb des Grenztetraeders T_1 von k, entsteht und daß die übrigen Teile von S und die Elementarflächenstücke $\mathfrak{g}$ und $\mathfrak{n}$ festbleiben. Dabei geht k in sich über. Daraufhin wird t_1' über das ausgebeulte Stück so deformiert, daß auf t_1' nur ein Minimum der zentrischen Distanz dicht bei dem jetzigen Endpunkt Q von k_1 entsteht, und dabei wird $\mathfrak{g}$ so mitgezogen, daß auf $\mathfrak{g}$ nur eine einzige Ausnahmeecke in der Nähe des bisherigen Minimums der zentrischen Distanz von t_1' entsteht. Diese liegt dann auf dem Rande einer inneren Haube von $\mathfrak{g}$ und kann daher so beseitigt werden, daß T_2' punktweise festbleibt. Nunmehr wird das aus $\mathfrak{k}_2$, $\mathfrak{s}_1$ und dem deformierten Knie $\mathfrak{k}_1$ bestehende Endstück von S längs k so verbeult, daß es wieder aus einem Knie, einem Schenkel und einem Kopf besteht und daß der auf der Grenzlinie von $\mathfrak{k}_1$ liegende Sehnenendpunkt auf dieselbe zentrale Strecke gerät, auf der jetzt Q liegt. Insgesamt wird also der

[23]) Vgl. [6] § 5, 5. Schritt.
[24]) Vgl. [6] § 6, 2. Schritt.
[25]) Vgl. [6] § 6, 1. Schritt.

9272-65

Kopf $\mathfrak{k}_1$ von einer zentralen Strecke von k über einen inneren Weg hinweg auf eine andere verschoben. Damit entsteht der Fall a.

Fall c. k_1 umfaßt einen inneren Weg von k und schneidet S.

Auf k_1 liegen also jetzt die beiden Schnittpunkte von k mit dem zweiten Kopf $\mathfrak{k}'$ von S. Liegen diese nicht auf derselben zentralen Strecke wie Q, so kann das Verfahren des Falles b angewandt werden. Liegen beide Schnittpunkte auf derselben zentralen Strecke wie Q, so betrachten wir den an $\mathfrak{k}'$ anstoßenden Schenkel. Führt dieser von $\mathfrak{k}'$ nach außen, so kann zunächst $\mathfrak{k}'$ längs k_1 über den inneren Weg hinweg auf die zweite zentrale Strecke geschoben werden und danach $\mathfrak{k}_1$ wie im Falle b.

Nehmen wir nun an, daß die Schnittpunkte von k mit dem zweiten Knie $\mathfrak{k}'$ auf k_1 auf derselben zentralen Strecke wie Q liegen und daß der an $\mathfrak{k}'$ anstoßende Schenkel $\mathfrak{s}'$ von $\mathfrak{k}'$ nach innen führt. Wie im Falle a erreichen wir zunächst, daß auf $\mathfrak{M}$ nur innere Knie liegen. Durch ein entsprechendes Verfahren können wir erreichen, daß auf $\mathfrak{M}$ alle inneren Knie innerhalb des Grenztetraeders T_1 von k gelangen. Auf $\mathfrak{M}$ liegt möglicherweise außer k_1 noch ein zweites Stück k_2 von k, das dann einen inneren Weg u_2 von k umfaßt. Wenn dies der Fall ist, so heben wir u_2 unter Verlängerung der anstoßenden zentralen Strecken ins Innere von T_1 ab, so daß u_2 auf ein Schartetraeder zwischen T_1 und der äußersten Ausnahmeecke der auf $\mathfrak{M}$ liegenden inneren Knie gelangt. Dadurch wird es möglich, den Kopf $\mathfrak{k}'$ längs k_1 ähnlich wie im Falle b auf die zweite zentrale Strecke zu verschieben. Daraufhin kann u_2 längs eines geeigneten Elementarflächenstückes auf T_1 zurückdeformiert werden und zwar so, daß wieder zentrale Strecken an u_2 anstoßen und daß S und k_1 von u_2 nicht getroffen werden, d. h. man weiche den Schenkeln von S und dem inneren Weg auf k_1 bei der Deformation aus. Die Gestalt von u_2 ändert sich dabei, sie kann jedoch wie im 1. Schritt von § 6 durch eine zentrische Abbildung, die außerhalb T_2' die Identität ist, wiederhergestellt werden. Nunmehr können die innerhalb T_1 liegenden Knie von S vermöge geeigneter Elementarflächenstücke so abgebogen werden, daß sie wieder außerhalb T_1 gelangen, wobei man den inneren Wegen von k offenbar ausweichen kann. Das Verfahren des Falles b ist nunmehr anwendbar.

Bei den vorangehenden Deformationen werden die Elementarflächenstücke $\mathfrak{e}_1$ und $\mathfrak{e}_2$ in Mitleidenschaft gezogen. Die Deformationen können jedoch stets so vorgenommen werden, daß $\mathfrak{e}_1$ und $\mathfrak{e}_2$ wieder die Forderungen des § 5 erfüllen. Wir haben erhalten, daß sich stets ein Knie von S beseitigen läßt, solange noch Knie vorhanden sind. Da bei keiner Reduktion neue Knie entstehen, können wir weiterhin annehmen, daß S keine Ausnahmeecken besitzt, also aus den beiden Köpfen und einem einzigen Schenkel besteht.

§ 8. Abschließende Reduktionen für $\mathfrak{f}_1$ und $\mathfrak{f}_2$

Wir beseitigen zunächst alle noch verbliebenen Ausnahmeecken von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ und nehmen anschließend im 4. Schritt eine Reduktion der Schnitte von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ mit den Grenztetraedern von k vor.

Nehmen wir an, daß etwa auf $\mathfrak{f}_1$ noch Ausnahmeecken vorhanden sind. Wegen § 5, 3. Schritt und § 6, 1. Schritt liegt jede Ausnahmeecke zwischen den Tetraedern durch die Grenzlinien der beiden Köpfe von S. Wir betrachten auf $\mathfrak{f}_1$ alle diejenigen Schnitte mit Schartetraedern, die von Ausnahmeecken und den singulären Extrema der zentrischen Distanz von r_1 ausgehen. Es existiert dann auf $\mathfrak{f}_1$ ein zu k punktfremdes Elementarflächenstück $\mathfrak{g}$, das entweder ein Zipfel ist oder von einem Querschnitt der angegebenen Art abgeschnitten wird und das keine weiteren solchen Schnitte enthält.

1. Schritt. *Reduktion, falls* $\mathfrak{g}$ *ein Zipfel ist.*

Es sei A die Ausnahmeecke auf dem Rande von $\mathfrak{g}$ und T das Schartetraeder durch A. Mit dem Rande von $\mathfrak{f}_1$ hat $\mathfrak{g}$ ein Stück t_1' gemein, das auf r_1 liegt, und es wird $\mathfrak{g}$ von t_1' und einem Querschnitt q von $\mathfrak{f}_1$ durch A berandet. t_1' führt mindestens einmal über einen Kopf von S, etwa über den Kopf $\mathfrak{k}_1$. Zur Fixierung der Sprechweise nehmen wir an, daß $\mathfrak{k}_1$ der innere der beiden Köpfe ist. Auf $\mathfrak{k}_1$ liegt ein reguläres Minimum A_1 der zentrischen Distanz von t_1'. Das Schartetraeder T' durch A_1 schneidet $\mathfrak{f}_1$ in einer geschlossenen Schnittlinie s durch A_1, die auf $\mathfrak{g}$ liegt. Dehnt man T' stetig, so geht s in einen Schnitt über, der auf S entspringt und endet und in q übergeht, wenn T' in T gedehnt ist. Wäre dies nämlich nicht der Fall, so könnte man die Dehnung über T hinaus fortsetzen, bis der aus s entstandene Schnitt in einen Querschnitt von $\mathfrak{g}$ übergeht, der in einem Extremum der zentrischen Distanz von t_1' auf dem äußeren Kopf von S entspringt, was der Wahl von $\mathfrak{g}$ widerspricht. $\mathfrak{g}$ hat also ähnliche Gestalt wie das im 3. Schritt von § 7 bestimmte, dort ebenfalls mit $\mathfrak{g}$ bezeichnete Elementarflächenstück, und ähnlich wie dort wird eine innerhalb T liegende Kugel $\mathfrak{M}$ bestimmt: q berandet auf T zusammen mit einem geeigneten Stück q' des Schnittes von S und T ein Elementarflächenstück $\mathfrak{n}$, das an $\mathfrak{K}$ von außen in q' anstößt und mit k nur einen Schnittpunkt gemein hat. q' und t_1' beranden auf S ein Elementarflächenstück $\mathfrak{m}$, das innerhalb T liegt und k in einem Sehnenendpunkt Q trifft. $\mathfrak{g}$, $\mathfrak{m}$ und $\mathfrak{n}$ beranden eine innerhalb T liegende Kugel $\mathfrak{M}$. Ihr Durchschnitt mit k besteht aus einem Stück k_1 von k, ihr Durchschnitt mit $\mathfrak{K}$ besteht aus $\mathfrak{m}$ nach Bestimmung von $\mathfrak{n}$. k_1 liegt entweder auf einer zentralen Strecke von k oder umfaßt einen inneren Weg. Wir betrachten hier nur den ersten Fall und zeigen im 3. Schritt, daß sich der zweite Fall auf den ersten zurückführen läßt.

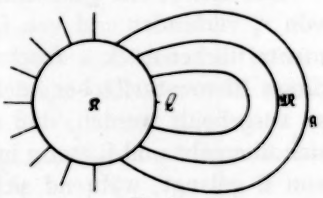


Fig. 11

Im betrachteten Fall liegt der Endpunkt Q von k_1 auf der Grenzlinie von $\mathfrak{k}_1$. Durchläuft man diese Grenzlinie, so trifft man abwechselnd auf die Ränder von $\mathfrak{f}_1$ und $\mathfrak{f}_2$. Die Schnitte des Tetraeders durch die Grenzlinie von $\mathfrak{k}_1$ mit $\mathfrak{f}_1$ und $\mathfrak{f}_2$ haben daher, soweit sie auf $\mathfrak{M}$ liegen, folgendes Aussehen (Fig. 11): Es sind Schnitte, die auf der Grenzlinie von $\mathfrak{k}_1$ entspringen und enden, Q trennt die Endpunkte jeder Schnittlinie auf demjenigen Stück der

Grenzlinie von $\mathfrak{k}_1$, das auf $\mathfrak{m}$ liegt, und der äußerste dieser Schnitte liegt auf $\mathfrak{g}$. Weitere Schnitte können nicht vorhanden sein, weil solche geschlossen sein und auf einer Haube von $\mathfrak{f}_1$ oder $\mathfrak{f}_2$ liegen müßten, wie man erkennt, wenn man das Schartetraeder durch die Grenzlinie von $\mathfrak{k}_1$ in geeignetem Sinne stetig dehnt. Dehnt man das Schartetraeder durch die Grenzlinie von $\mathfrak{k}_1$ stetig nach T, so entsteht aus Q ein Querschnitt von $\mathfrak{f}_1$ oder $\mathfrak{f}_2$, etwa von $\mathfrak{f}_2$, während die übrigen auf $\mathfrak{M}$ liegenden Schnitte ihre Gestalt behalten. Von zwei auf dem Tetraeder benachbarten Schnitten liegt nämlich einer auf $\mathfrak{f}_1$, der andere auf $\mathfrak{f}_2$, so daß eine Veränderung des Schnittbildes nur dann möglich wäre, wenn noch Hauben vorhanden wären. Zugleich bemerkt man, daß von den vier Schnitten, die auf T von A ausgehen, nur die beiden, die zusammen q ausmachen, auf $\mathfrak{n}$ liegen.

Auf $\mathfrak{n}$ liegt ein Querschnitt, der den Endpunkt von k_1 mit einem Punkte von r_2 verbindet und von $\mathfrak{f}_2$ ein ausnahmeeckenfreies, auf $\mathfrak{M}$ liegendes Elementarflächenstück $\mathfrak{a}$ abschneidet, das Q als innersten Punkt besitzt. Längs dieses Elementarflächenstückes kann $\mathfrak{k}_1$ und das anschließende Schenkelstück so ausgebeult werden, daß die Grenzlinie von $\mathfrak{k}_1$ außerhalb T gelangt, k in sich übergeht und $\mathfrak{f}_2$ stetig in sich zusammengezogen wird, so daß $\mathfrak{a}$ ins Innere von $\mathfrak{R}$ gelangt, während sich $\mathfrak{f}_1$ und $\mathfrak{f}_2$ sonst nicht ändern. Nach dem Verfahren des 3. Schrittes von § 5 kann nun $\mathfrak{k}_1$ so zusammengezogen werden, daß $\mathfrak{k}_1$ außerhalb T zu liegen kommt. $\mathfrak{g}$ geht dabei in eine Haube über, und die Ausnahmeecke A kann beseitigt werden. Bei dem Verbeulen von $\mathfrak{k}_1$ ist so vorzugehen, daß das anstoßende der beiden Elementarflächenstücke $\mathfrak{e}_1$, $\mathfrak{e}_2$ auch anschließend den Forderungen des § 5 genügt.

2. Schritt. *Reduktion, falls* $\mathfrak{g}$ *von einem Querschnitt abgeschnitten wird, der von einem singulären Extremum der zentrischen Distanz von* r_1 *ausgeht.*

Sei t_1 das Stück von r_1, das $\mathfrak{g}$ und $\mathfrak{f}_1$ gemein haben. Da r_1 abwechselnd über die beiden Köpfe von S geht, liegt auf t_1 ein Extremum A_1 der zentrischen Distanz, das nicht Endpunkt ist. Zur Fixierung der Sprechweise nehmen wir wieder an, daß A_1 auf dem inneren der beiden Köpfe von S liegt.

Es sei jetzt T ein Schartetraeder, das den Schenkel von S schneidet, und zwar so dicht an der Grenzlinie des äußeren Kopfes, daß auf T und außerhalb T keine Ausnahmeecken von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ liegen. T schneidet $\mathfrak{g}$ in einem Querschnitt q, und es berandet q zusammen mit einem Stück t_1' von t_1, das A_1 enthält, ein Elementarflächenstück auf $\mathfrak{g}$, dem wir von jetzt an die Bezeichnung $\mathfrak{g}$ erteilen. Nun kann die Kugel $\mathfrak{M}$ wie im 1. Schritt bestimmt werden, wobei die Bezeichnungen entsprechend zu wählen sind. Der Durchschnitt von $\mathfrak{M}$ und k besteht wieder aus einem Stück k_1 von k, wobei k_1 entweder auf einer zentralen Strecke von k liegt oder einen inneren Weg umfaßt. Der zweite Fall wird durch den 3. Schritt auf den ersten zurückgeführt, so daß wir hier nur diesen zu betrachten haben. Der Vorgang entspricht nun dem 1. Schritt. Wie dort kann das Elementarflächenstück $\mathfrak{a}$ bestimmt und S längs $\mathfrak{a}$ so verbeult werden, daß die Grenzlinie von $\mathfrak{k}_1$ außerhalb T gelangt aber noch innerhalb des Schartetraeders durch die Grenzlinie des

äußeren Kopfes liegt, und es läßt sich dann $\mathfrak{k}_1$ außerhalb T bringen. Dann hat keine Ausnahmeecke von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ dieselbe zentrische Distanz wie ein Punkt von S, und es verschwinden alle, wenn man Hauben auf $\mathfrak{f}_1$ und $\mathfrak{f}_2$ beseitigt, weil von einer Ausnahmeecke nicht alle Schnitte mit dem zugehörigen Schartetraeder zu k führen können (vgl. § 6).

3. Schritt. *Reduktion, falls k_1 in den vorangehenden Schritten einen inneren Weg von k umfaßt.*

Sei u_1 dieser innere Weg, die übrigen Bezeichnungen seien dieselben wie im Vorangehenden. Die Reduktion besteht darin, daß $\mathfrak{k}_1$ über u_1 hinweggezogen wird, worauf der bisher betrachtete Sachverhalt des 1. oder 2. Schrittes entsteht.

Der auf $\mathfrak{k}_1$ liegende Endpunkt Q von k_1 ist jetzt der innerste Punkt von S. An $\mathfrak{k}_1$ stößt eines der beiden Elementarflächenstücke $\mathfrak{e}_1$, $\mathfrak{e}_2$ an, etwa $\mathfrak{e}_1$. Die Schnitte von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ mit dem Schartetraeder T' durch die Grenzlinie von $\mathfrak{k}_1$ haben, soweit sie auf $\mathfrak{M}$ liegen, das Aussehen von Fig. 11, wobei jetzt nur an Stelle des dort mit Q bezeichneten Punktes ein Querschnitt p etwa von $\mathfrak{f}_2$ tritt, der einen Punkt von k_1 mit einem Punkt von r_2 verbindet. p schneidet von $\mathfrak{f}_2$ ein Elementarflächenstück $\mathfrak{b}$ ab, das auf $\mathfrak{M}$ liegt und das von p, einem Stück von k_1 und einem in Q entspringenden Stück von r_2 berandet wird (Fig. 12). Längs $\mathfrak{b}$ kann $\mathfrak{k}_1$ und das anstoßende Schenkelstück so ausgebeult werden, daß k in sich übergeht, daß auf S keine Ausnahmeecke entsteht und daß $\mathfrak{b}$ ins Innere von $\mathfrak{R}$ gelangt, wobei $\mathfrak{f}_2$ in sich zusammengezogen wird, während sich $\mathfrak{f}_1$ und $\mathfrak{f}_2$ sonst nicht ändern. $\mathfrak{e}_1$ ist dabei so mitzunehmen, daß auf $\mathfrak{e}_1$ keine Ausnahmeecken entstehen, daß $\mathfrak{e}_1$ an u_1 vom Inneren des Grenztetraeders T_1 her anstößt und daß der innerste Punkt des deformierten Kopfes $\mathfrak{k}_1$ auf dem Rande von $\mathfrak{e}_1$ liegt. Bei der Deformation gelangt Q außerhalb T'. Die innerhalb T liegenden Stücke von r_1 und r_2 (T hat die Bedeutung wie im 1. oder 2. Schritt) werden nun auf dem ausgebeulten Stück so deformiert, daß wieder jeweils nur ein Minimum der zentrischen Distanz vorhanden ist und daß dieses dicht bei dem jetzigen Sehnenendpunkt Q außerhalb T' liegt. Die Deformation geschieht wie in § 6, 2. Schritt, Fall a. Die anstoßenden Teile von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ sind mitzuziehen, wobei neue Ausnahmeecken entstehen, die jedoch innerhalb T' liegen. Nun wird $\mathfrak{k}_1$ weiter längs $\mathfrak{e}_1$ und eines Teiles von $\mathfrak{b}$ so verbeult, daß der Schnittpunkt Q' mit k, der anfangs auf der Grenzlinie von $\mathfrak{k}_1$ lag, auf dieselbe zentrale Strecke wie Q gelangt, zum Minimum der zentrischen Distanz auf S wird und außerhalb T' zu liegen kommt. Dies muß so geschehen, daß auf S keine Ausnahmeecken entstehen und k in sich übergeht. Q' ist anfangs Minimum der zentrischen

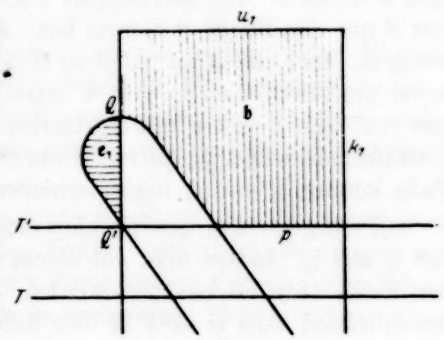


Fig. 12

Distanz auf $\mathfrak{f}_1$ oder $\mathfrak{f}_2$. Das anstoßende Stück von $\mathfrak{f}_1$ oder $\mathfrak{f}_2$ kann bei der Deformation so mitgezogen werden, daß es außerhalb des Grenztetraeders T_1 verbleibt und keine Ausnahmeecken entstehen, weil e_1 und das betreffende Stück von $\mathfrak{f}_1$ oder $\mathfrak{f}_2$ nur den Punkt Q' gemein haben. S liegt nunmehr außerhalb T', und die bei der Deformation entstandenen Ausnahmeecken verschwinden, wenn Hauben auf $\mathfrak{f}_1$ und $\mathfrak{f}_2$ wieder beseitigt werden. Der gewünschte Sachverhalt ist damit erreicht.

4. Schritt. *Korrektur der Schnitte durch die regulären Extrema.*

Sei A ein solches Extremum etwa auf r_1, das nicht Sehnenendpunkt ist. Das Schartetraeder T durch A schneidet $\mathfrak{f}_1$ in einer geschlossenen Schnittlinie s durch A, und es berandet s auf T ein Elementarflächenstück $\mathfrak{n}$, das mit S nur den Punkt A gemein hat. Auf $\mathfrak{f}_1$ berandet s ein Elementarflächenstück $\mathfrak{m}$. Sei $\mathfrak{k}_1$ der Kopf, auf dem A liegt. $\mathfrak{n}$ trifft die Sehne von $\mathfrak{k}_1$ nicht, und $\mathfrak{m}$ ist punktfremd zu k. Auf $\mathfrak{n}$ liegen daher zwei oder keine Schnittpunkte mit k. Wenn k von $\mathfrak{n}$ nicht getroffen wird, so beranden $\mathfrak{m}$ und $\mathfrak{n}$ eine zu k punktfremde Kugel $\mathfrak{M}$, die mit S nur den Punkt A gemein hat, und in diesem Falle können $\mathfrak{f}_1$ und $\mathfrak{f}_2$ folgendermaßen deformiert werden:

Auf $\mathfrak{n}$ liegen außer dem Rande s möglicherweise noch geschlossene Schnitte mit $\mathfrak{f}_1$ und $\mathfrak{f}_2$. Indem man mit einem auf $\mathfrak{n}$ innersten beginnt, können diese der Reihe nach so beseitigt werden, daß keine Ausnahmeecken entstehen. Anschließend wird $\mathfrak{m}$ über $\mathfrak{M}$ in $\mathfrak{n}$ deformiert und so von $\mathfrak{n}$ abgehoben, daß k, S und A festbleiben, A zum Extremum der zentrischen Distanz auf $\mathfrak{f}_1$ wird und keine Ausnahmeecken entstehen. Das Verfahren soll so oft wie möglich auf $\mathfrak{f}_1$ und $\mathfrak{f}_2$ angewandt werden.

§ 9. Nachweis, daß $\mathfrak{f}_1$ und $\mathfrak{f}_2$ eines der beiden Grenztetraeder von k nicht mehr schneiden und in je einem Stück von k berühren

1. Schritt. *Nachweis, daß auf r_1 und r_2 insgesamt α' reguläre Extrema der zentrischen Distanz vorhanden sind.*

Auf r_1 und r_2 liegen je $\alpha'-1$ Extrema der zentrischen Distanz, die nicht Sehnenendpunkte sind. Unter diesen betrachten wir die singulären. Von jedem solchen Extremum gehen zwei Querschnitte von $\mathfrak{f}_1$ oder $\mathfrak{f}_2$ auf dem zugehörigen Schartetraeder aus, die $\mathfrak{f}_1$ bzw. $\mathfrak{f}_2$ jeweils in drei Elementarflächenstücke zerlegen, wobei eines von beiden Querschnitten und einem Stück des Randes von $\mathfrak{f}_1$ bzw. $\mathfrak{f}_2$ berandet wird, also „zipfelartige" Gestalt hat, während die anderen beiden von je einem Querschnitt und einem Stück des Randes von $\mathfrak{f}_1$ bzw. $\mathfrak{f}_2$ berandet werden. Auf $\mathfrak{f}_1$ und $\mathfrak{f}_2$ betrachten wir alle diese Querschnittpaare zugleich, und es mögen γ_1 bzw. γ_2 solche Paare auf $\mathfrak{f}_1$ bzw. $\mathfrak{f}_2$ liegen. Von $\mathfrak{f}_1$ bzw. $\mathfrak{f}_2$ werden dann γ_1+2 bzw. γ_2+2 Elementarflächenstücke abgeschnitten, die mit den Querschnitten nur Randpunkte gemein haben, wie durch vollständige Induktion nach γ_1, γ_2 folgt. Sei $\mathfrak{g}$ ein solches Elementarflächenstück, etwa auf $\mathfrak{f}_1$. Für die Lage von $\mathfrak{g}$ sind drei Fälle möglich:

a) $\mathfrak{g}$ trifft k nicht,

b) $\mathfrak{g}$ hat mit k ein Stück gemein, das einen Sehnenendpunkt als Endpunkt besitzt,

c) $\mathfrak{g}$ hat zipfelartige Gestalt und mit r_1 nur einen Punkt, die „Spitze des Zipfels", gemein.

Im Falle a liegen auf dem Rande von $\mathfrak{g}$ genau zwei Extrema der zentrischen Distanz von r_1, nämlich ein singuläres, von dem zwei Querschnitte von $\mathfrak{f}_1$ mit dem zugehörigen Schartetraeder ausgehen, von denen einer oder beide zum Rande von $\mathfrak{g}$ gehören, während das andere Extremum regulär ist. Man erkennt dies, wenn man das Schartetraeder durch das singuläre Extremum in geeignetem Sinne stetig dehnt, so daß auf $\mathfrak{g}$ ein Schnitt entsteht, und diesen Schnitt verfolgt.

Im Falle b erhält man entsprechend zwei Extrema der zentrischen Distanz von r_1, wobei das zweite hier ein Sehnenendpunkt ist. Ist dieses ein reguläres Extremum, so hat $\mathfrak{g}$ mit k nur ein Stück einer zentralen Strecke gemein, andernfalls liegt auf dem Rande von $\mathfrak{g}$ ein innerer oder ein äußerer Weg von k und nur einer.

Im Falle c liegt auf dem Rande von $\mathfrak{g}$ nur ein Extremum der zentrischen Distanz von r_1, nämlich die Spitze des Zipfels, außerdem aber genau ein innerer oder äußerer Weg von k, wie sich durch stetige Dehnung des Schartetraeders durch die Spitze des Zipfels in geeignetem Sinne ergibt.

Berücksichtigt man, daß auf k zwei innere und zwei äußere Wege liegen sowie vier Sehnenendpunkte, von denen zwei reguläre Extrema sind, so bemerkt man, daß für $\mathfrak{f}_1$ und $\mathfrak{f}_2$ zusammen der Fall b genau viermal, der Fall c genau zweimal, der Fall a also $(\gamma_1+\gamma_2-2)$-mal eintritt.

Betrachten wir nun ein Schartetraeder durch ein reguläres Extremum der zentrischen Distanz auf r_1 oder r_2, das nicht Sehnenendpunkt ist. Wenn man das Tetraeder in geeignetem Sinne stetig dehnt, erkennt man, daß das Extremum auf einem Elementarflächenstück $\mathfrak{g}$ von $\mathfrak{f}_1$ bzw. $\mathfrak{f}_2$ liegt, das zum Fall a gehört. Zugleich ergibt sich, daß γ_1 und γ_2 von null verschieden sind, was oben stillschweigend vorausgesetzt wurde. Wir erhalten, daß auf r_1 und r_2 zusammen $2\gamma_1+2\gamma_2-2$ Extrema der zentrischen Distanz liegen, die nicht Sehnenendpunkte sind. Also ist $\gamma_1+\gamma_2=\alpha'$, und dies ist auch die Anzahl der regulären Extrema der zentrischen Distanz auf r_1 und r_2, weil genau zwei Sehnenendpunkte solche Extrema darstellen.

2. Schritt. *Verhalten der Schnitte von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ mit einem stetig gedehnten Schartetraeder.*

Es sei T ein Schartetraeder, das den Schenkel von S schneidet. Es liegt also einer der beiden Köpfe von S innerhalb T, der andere außerhalb. Der Schnitt von S und T sei s. Durchläuft man s einmal, so gelangt man abwechselnd zu Schnitten von r_1 und r_2 mit T, und zwar je α'-mal. Der Durchschnitt von T und $\mathfrak{f}_1$, $\mathfrak{f}_2$ besteht daher (Fig. 13) aus vier Querschnitten q_1, q_2, q_3, q_4, die von s nach je einem der vier Schnittpunkte von k und T führen, und

Querschnitten, die auf s entspringen und enden und dabei jeweils genau einen der Querschnitte q_1, q_2, q_3, q_4 umfassen. Weitere Schnitte von T mit $\mathfrak{f}_1$ und $\mathfrak{f}_2$ sind nicht vorhanden, weil es auf $\mathfrak{f}_1$ und $\mathfrak{f}_2$ keine Ausnahmeecken mehr gibt. Der Querschnitt q_i $(i=1, 2, 3, 4)$ werde von n_i Querschnitten umfaßt, wobei die Numerierung der q_i entsprechend Fig. 13 gewählt sei. Dann ist

$$n_1 + n_2 + n_3 + n_4 = \alpha' - 2. \tag{1}$$

Zieht man T stetig auf seinen Mittelpunkt zusammen, so ändert sich das Schnittbild erst dann, wenn T durch die Grenzlinie des inneren Kopfes von S geht, und es verschwindet dort einer der Querschnitte q_i. Die Bezeichnung sei so gewählt, daß dies q_1 ist. Wird T weiter zusammengezogen, so schrumpfen die n_1 Schnitte, die vorher q_1 umfaßten, der Reihe nach auf einen Punkt zusammen wegen des 4. Schrittes von § 9. Man erhält so zunächst n_1+1 reguläre Extrema der zentrischen Distanz von r_1 und r_2. Nehmen wir an, daß $n_2 \leq n_4$ ist, so verschmelzen bei weiterem Zusammenziehen von T nun n_2 Querschnitte, die q_2 umfassen, mit solchen, die q_4 umfassen, und ein solches Verschmelzen von Querschnitten tritt auch weiterhin solange ein, bis ein Schnitt entsteht, der zwei Punkte von k verbindet. Danach entsteht eine gewisse Anzahl δ geschlossener Schnitte, bis schließlich T durch den innersten Punkt von S geht und zwei Querschnitte auf $\mathfrak{f}_1$ und $\mathfrak{f}_2$ vorhanden sind, die je zwei Punkte von k verbinden, und δ geschlossene Schnitte, welche die beiden Querschnitte auf T trennen und bei weiterem Zusammenziehen in δ geschlossene Schnitte von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ mit dem inneren Grenztetraeder T_1 übergehen. Das Verschmelzen der Schnitte ist in Fig. 13 gestrichelt angedeutet.

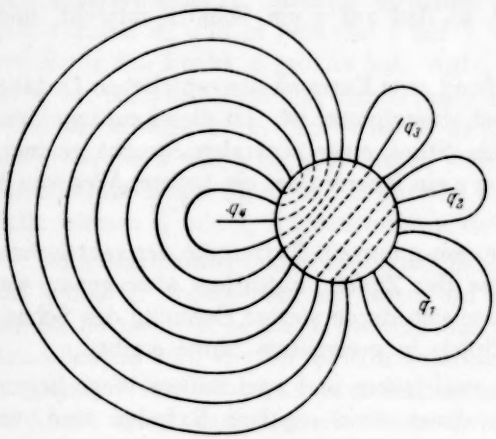


Fig. 13

Um δ zu bestimmen, betrachten wir T in derjenigen Lage, in der q_2 mit einem anderen Querschnitt verschmilzt, und lassen q_2 fort. Auf dem von T aus $\mathfrak{K}$ ausgeschnittenen Elementarflächenstück $\mathfrak{h}$ verbinden wir die Schnitte von T mit $\mathfrak{f}_1$ und $\mathfrak{f}_2$ entsprechend dem weiteren Verschmelzen (Fig. 14). Wir erhalten dann zwei nicht geschlossene Wege und δ geschlossene, welche die beiden ersten trennen. Auf dem Rande s von $\mathfrak{h}$ liegen zwei Endpunkte und $2n_3+2n_4$ Schnittpunkte mit diesen Wegen. Die beiden auf s liegenden Endpunkte können über $\mathfrak{h}$ hinweg durch einen Weg w_1 so verbunden werden, daß n_3+n_4 Schnittpunkte entstehen. Entsprechend können die beiden anderen Endpunkte der nicht geschlossenen Wege durch einen Weg w_2 außerhalb $\mathfrak{h}$ verbunden werden. Das entstehende Bild kann als Projektion einer Normal-

form mit zwei Brücken aufgefaßt werden, zu der zusätzlich δ geschlossene Wege gezeichnet sind, welche die beiden äußeren Wege trennen und die Brücken abwechselnd unterkreuzen[26]). Zählt man die geschlossenen Wege mit, so liest man n_3+n_4+1 als die der Torsion entsprechende Zahl und n_4-n_2 oder n_2+n_3+1 als Betrag der der Kreuzungsklasse entsprechenden Zahl ab. Nach einer Bemerkung in § 3 erhält man für die Anzahl der geschlossenen Wege jedenfalls:

$$(2) \qquad \delta = (n_4 - n_2,\ n_3 + n_4 + 1) - 1 = (n_2 + n_3 + 1,\ n_3 + n_4 + 1) - 1,$$

wobei die Klammer den größten gemeinsamen Teiler bezeichnet.

Wir erhalten: Die Anzahl der Schnitte von T_1 mit $\mathfrak{f}_1$ und $\mathfrak{f}_2$ ist $(n_2+n_3+1,\ n_3+n_4+1)-1$, wenn die inneren Wege von k, in denen T_1 von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ berührt wird, nicht als Schnitte gezählt werden. Zugleich hat sich ergeben, daß auf dem inneren Kopf von S insgesamt $n_1+(n_2+n_3+1,\ n_3+n_4+1)$ reguläre Extrema der zentrischen Distanz von r_1 und r_2 liegen. Diese Aussagen wurden unter der Annahme $n_2 \leq n_4$ hergeleitet, gelten aber wegen der Symmetrie in n_2 und n_4 allgemein.

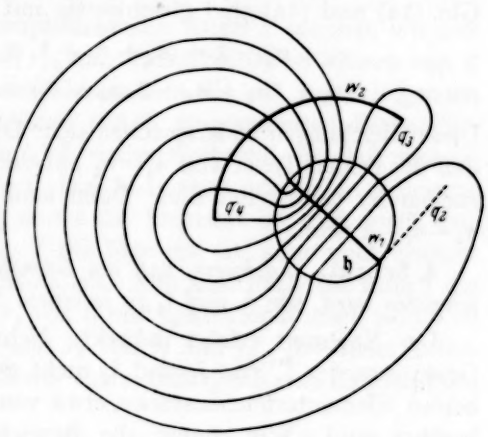


Fig. 14

3. Schritt. *Nachweis, daß eines der beiden Grenztetraeder von* $\mathfrak{f}_1$ *und* $\mathfrak{f}_2$ *nicht geschnitten wird.*

Das im Vorangehenden benutzte Tetraeder T betrachten wir wieder in der Ausgangslage (Fig. 13) und dehnen es stetig bis zur Grenzlinie des äußeren Kopfes von S. Es verschwindet dann einer der Querschnitte q_i. Dies kann nicht q_1 sein, da sonst r_1 oder r_2 mit den Köpfen nur die Endpunkte gemein hätte und somit $\alpha'=1$ wäre, was ausgeschlossen wurde. Es sind zwei Fälle zu unterscheiden, je nachdem q_3 oder einer der beiden Querschnitte q_2, q_4 verschwindet.

Fall a. Auf der Grenzlinie des äußeren Kopfes verschwindet q_2 *oder* q_4.

Wir können annehmen, daß q_2 verschwindet, der andere Fall entsteht durch Umbezeichnung. Entsprechend dem 2. Schritt ergeben sich auf dem äußeren Kopf $n_2+(n_3+n_4+1,\ n_1+n_4+1)$ reguläre Extrema der zentrischen Distanz von r_1 und r_2. Wegen der beiden vorangehenden Schritte und Gl. (1) muß sein

$$(3) \qquad n_3 + n_4 + 2 = (n_2 + n_3 + 1,\ n_3 + n_4 + 1) + (n_1 + n_4 + 1,\ n_3 + n_4 + 1).$$

Diese Beziehung kann aber nur bestehen, wenn einer der beiden größten gemeinsamen Teiler gleich 1, der andere gleich n_3+n_4+1 ist. Hieraus folgt

[26]) Der Sachverhalt läßt sich leicht verfolgen, wenn man das Schartetraeder durch den innersten Punkt von S stetig dehnt.

wegen Gl. (2) und der entsprechenden Gleichung für das äußere Grenztetraeder, daß eines der beiden Grenztetraeder von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ nicht geschnitten wird.

Fall b. Auf der Grenzlinie des äußeren Kopfes verschwindet q_3.

Dieser Fall ist nicht möglich. Entsprechend Gl. (3) erhielte man

(3a) $$n_2 + n_4 + 2 = (n_2 + n_3 + 1, n_3 + n_4 + 1) + (n_1 + n_2 + 1, n_1 + n_4 + 1).$$

Nach Definition von β' erhielte man ferner, daß unter der unwesentlichen Annahme $n_2 \leq n_4$ entweder $|\beta'| = n_1 + 2n_2 + n_3 + 2$ oder $\alpha' - |\beta'| = n_1 + 2n_2 + n_3 + 2$ wäre. Wegen der Teilerfremdheit von α' und β' und wegen Gl. (1) müßte jedenfalls sein

(4a) $$(n_1 + 2n_2 + n_3 + 2, n_1 + n_2 + n_3 + n_4 + 2) = 1.$$

Gln. (3a) und (4a) sind gleichwertig mit

$$n_2 + n_4 + 2 = (n_2 + n_3 + 1, n_4 - n_2) + (n_1 + n_2 + 1, n_4 - n_2)$$
$$(n_1 + 2n_2 + n_3 + 2, n_4 - n_2) = 1.$$

Diese Gleichungen widersprechen sich: Die zweite besagt, daß $n_4 \neq n_2$ ist und daß die beiden Teiler von $n_4 - n_2$ auf der rechten Seite der ersten Gleichung zueinander teilerfremd sind. Dann kann aber ihre Summe nicht größer als $n_4 - n_2 + 1$ sein.

4. Schritt. *Nachweis, daß ein Grenztetraeder, das von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ nicht geschnitten wird, von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ in je einem Stück von k berührt wird.*

Der Nachweis erfolgt indirekt. Nehmen wir etwa an, daß das äußere Grenztetraeder T_2 von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ nicht geschnitten wird und von einem der beiden Elementarflächenstücke, etwa von $\mathfrak{f}_1$, in beiden äußeren Wegen von k berührt wird. Wir können die Bezeichnungen des Falles a im 3. Schritt zugrunde legen. Es gelten also Gln. (1), (2) und (3). Betrachten wir wieder den Schnitt mit einem Schartetraeder T zwischen den beiden Köpfen von S (Fig. 13), so müßten q_1, q_3 und q_4 auf $\mathfrak{f}_1$ und q_2 auf $\mathfrak{f}_2$ liegen. Da auf T die Schnitte mit $\mathfrak{f}_1$ und $\mathfrak{f}_2$ abwechseln, müßten $n_1 + n_4$ und $n_3 + n_4$ ungerade sein, was der Aussage $(n_1 + n_4 + 1, n_3 + n_4 + 1) = 1$, daß also T_2 nicht geschnitten wird, widerspricht.

§ 10. Abschluß des Beweises

$\mathfrak{f}_1$ und $\mathfrak{f}_2$ schneiden eines der beiden Grenztetraeder T_1, T_2 von k nicht. Nehmen wir zunächst an, daß dies für das äußere Grenztetraeder T_2 der Fall ist. Dann gibt es ein Tetraeder T zwischen T_1 und T_2, das S im Inneren enthält und $\mathfrak{f}_1$ und $\mathfrak{f}_2$ in nur je einem Querschnitt schneidet, wobei dieser Querschnitt $\mathfrak{f}_1$ bzw. $\mathfrak{f}_2$ in ein außerhalb T liegendes und ein innerhalb T liegendes Elementarflächenstück zerlegt.

Längs der innerhalb T liegenden Stücke von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ beulen wir $\mathfrak{R}$ so aus, daß der Durchschnitt von T und $\mathfrak{R}$ aus zwei Elementarflächenstücken $\mathfrak{h}_1$ und $\mathfrak{h}_2$ auf S besteht, wobei $\mathfrak{h}_1$ bzw. $\mathfrak{h}_2$ den bisherigen Schnitt von $\mathfrak{f}_1$ bzw. $\mathfrak{f}_2$ mit T umfaßt. Das Ausbeulen soll so geschehen, daß dabei k in sich übergeht und $\mathfrak{f}_1$ und $\mathfrak{f}_2$ auf die außerhalb T liegenden Stücke zusammengezogen werden. $\mathfrak{e}_1$ und $\mathfrak{e}_2$ verlieren dabei ihre bisherige Gestalt, jedoch spielt diese

keine Rolle mehr. Nun bildet $\mathfrak{R}$ zusammen mit der Kugel, die das „Äußere" von T bezüglich der S^3 ausmacht, einen Vollring, für den $\mathfrak{h}_1$ und $\mathfrak{h}_2$ Meridianflächen sind. Dieser Vollring ist unverknotet. Man erhält nämlich leicht eine Seele, die eine verallgemeinerte Brückendarstellung mit einer Brücke ist, wenn man vor dem Ausbeulen die Sehnen von $\mathfrak{R}$ innerhalb $\mathfrak{R}$ durch einen zentrisch monotonen Streckenzug verbindet und diesen vermöge zweier geeigneter Stücke der Brücken von k und eines Weges auf T_2 zu einer Knotenlinie ergänzt. Daraus folgt[27]), daß das Komplement von $\mathfrak{R}$ bezüglich des Inneren von T ein unverknoteter Vollring ist, für den die Ränder von $\mathfrak{h}_1$ und $\mathfrak{h}_2$ Breitenkreise sind, und es läßt sich $\mathfrak{R}$ bei festem k und festem $\mathfrak{f}_1$ und $\mathfrak{f}_2$ in das Innere von T überführen[28]).

In die Sehnen von $\mathfrak{R}$ und der komplementären Kugel $\mathfrak{L}$ spannen wir nun Elementarflächenstücke $\bar{\mathfrak{e}}_1, \bar{\mathfrak{e}}_2$ bzw. $\bar{\mathfrak{f}}_1, \bar{\mathfrak{f}}_2$ ein, indem wir diese Sehnen von Z aus auf T projizieren. Nach Hilfssatz 3 lassen sich die auf $S = T$ liegenden Randstücke s_1 und s_2 der in die Sehnen von $\mathfrak{R}$ eingespannten Elementarflächenstücke $\mathfrak{e}_1, \mathfrak{e}_2$ in diejenigen von $\bar{\mathfrak{e}}_1, \bar{\mathfrak{e}}_2$ dadurch überführen, daß man sie auf T kombinatorisch isotop deformiert. Dies kann auch durch eine zentrische Abbildung geschehen, die auf T_1, T_2 und k die Identität ist, und durch eine s-Abbildung, die auf k und außerhalb T die Identität ist, kann noch erreicht werden, daß $\mathfrak{e}_1, \mathfrak{e}_2$ mit $\bar{\mathfrak{e}}_1, \bar{\mathfrak{e}}_2$ zusammenfallen. Nun sind nach Hilfssatz 3 die auf T liegenden Randstücke r_1 und r_2 von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ kombinatorisch isotop zu den entsprechenden Randstücken $\bar{r}_1$ und $\bar{r}_2$ von $\bar{\mathfrak{f}}_1$ und $\bar{\mathfrak{f}}_2$. Bei dieser Deformation ändert sich das Kreuzungsschema von r_1 und r_2 mit den Randstücken s_1 und s_2 von $\mathfrak{e}_1$ und $\mathfrak{e}_2$ nicht:

Durch Deformation von s_1, s_2, r_1, r_2, die s_1 und s_2 in sich überführt, kann zunächst erreicht werden, daß der Durchschnitt von r_1 und r_2 mit $\bar{r}_1$ und $\bar{r}_2$ nur aus den Endpunkten und solchen Schnittpunkten besteht, die nicht auf s_1 oder s_2 liegen. Die weitere Deformation erfolgt wie im 1. Schritt des Beweises von Satz 2 (§ 2). Es erfolgen jeweils Deformationen über Elementarflächenstücke $\mathfrak{g}$, die von je einem Stück w' auf $\bar{r}_1$ oder $\bar{r}_2$ und einem Stück v' auf r_1 oder r_2 berandet werden, wobei w' außer seinen Endpunkten keine Schnittpunkte mit r_1 und r_2 besitzt und $\mathfrak{g}$ keinen der Endpunkte von r_1 und r_2 im Inneren enthält. s_1 und s_2 müssen auf einem solchen Elementarflächenstück $\mathfrak{g}$ Querschnitte bilden, die jeweils einen Punkt von w' mit einem Punkt von v' verbinden. Andernfalls würde w' oder v' zweimal hintereinander von s_1 oder s_2 geschnitten. Dies ist aber nicht möglich, weil s_1, s_2 mit r_1, r_2 und mit $\bar{r}_1, \bar{r}_2$ Kreuzungsschemata von Normalformen bilden, wie aus der Wahl der zugehörigen Elementarflächenstücke auf $\mathfrak{R}$ und $\mathfrak{L}$ folgt. Jede Deformation läßt sich also durch eine s-Abbildung von T auf sich bewirken, die s_1 und s_2 in sich überführt, und dies kann auch durch eine zentrische Abbildung geschehen, die auf k, T_1 und T_2 die Identität ist.

Wir haben erreicht, daß r_1 und r_2 mit $\bar{r}_1$ und $\bar{r}_2$ zusammenfallen. Unter der Annahme, daß T_2 von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ nicht geschnitten wird, ergibt sich also

[27]) [7] § 7, Satz 2.
[28]) [7] § 8.

$\alpha = \alpha'$ und $\beta = \beta'$. Wenn T_1 von $\mathfrak{f}_1$ und $\mathfrak{f}_2$ nicht geschnitten wird, läßt sich entsprechend vorgehen, wobei nur $\mathfrak{R}$ in das Äußere eines geeigneten Tetraeders überzuführen ist. Nach § 3 ergibt sich entsprechend $\alpha = \alpha'$ und $\beta\beta' \equiv 1$ modulo 2α. Der Beweis des Satzes ist damit beendet.

Anhang: Knoten der Alexander-Briggsschen Tabelle, für die Darstellungen mit zwei Brücken gefunden wurden

Alle Knoten der ALEXANDER-BRIGGSschen Tabelle[29]) besitzen Brückendarstellungen mit drei Brücken. Dies wurde mit Hilfe der Tatsache festgestellt, daß sich in einer Brückendarstellung eines Knotens eine Brücke stets dann beseitigen läßt, wenn sie von einem der beiden anstoßenden äußeren Wege nicht unterkreuzt wird. Brückendarstellungen mit zwei Brücken wurden für die folgenden Knoten gefunden:

Bezeichnung in Tabelle	Normalform	Bezeichnung in Tabelle	Normalform	Bezeichnung in Tabelle	Normalform
3,1	(3,1)	8,3	(17,13)	9,7	(29,9)
4,1	(5,3)	8,4	(19,5)	9,8	(31,11)
5,1	(5,1)	8,6	(23,7)	9,9	(31,7)
5,2	(7,3)	8,7	(23,5)	9,10	(33,23)
6,1	(9,5)	8,8	(25,9)	9,11	(33,7)
6,2	(11,3)	8,9	(25,7)	9,12	(35,13)
6,3	(13,5)	8,11	(27,17)	9,13	(37,11)
7,1	(7,1)	8,12	(29,17)	9,14	(37,23)
7,2	(11,5)	8,13	(29,11)	9,15	(39,17)
7,3	(13,3)	8,14	(31,13)	9,17	(39,25)
7,4	(15,11)	9,1	(9,1)	9,18	(41,17)
7,5	(17,5)	9,2	(15,7)	9,20	(41,11)
7,6	(19,7)	9,3	(19,3)	9,21	(43,25)
7,7	(21,13)	9,4	(21,5)	9,27	(49,19)
8,1	(13,7)	9,5	(23,17)	9,31	(55,21)
8,2	(17,3)	9,6	(27,5)		

Nach dem Korollar zu Satz 4 sind die Knoten 4,1; 6,3; 8,3; 8,9; 8,12 der Tabelle amphicheiral, was bereits von GOERITZ [2] bemerkt wurde.

Literatur

[1] BANKWITZ, C., u. H. G. SCHUMANN: Über Viergeflechte. Abh. Math. Sem. Hamburg **10**, 263—284 (1934). — [2] GOERITZ, L.: Bemerkungen zur Knotentheorie. Abh. Math. Sem. Hamburg **10**, 201—210 (1934). — [3] GRAEUB, W.: Die semilinearen Abbildungen. Sitzgsber. Akad. Wiss. Heidelberg, math.-nat. Kl., 4. Abh. **1950**. — [4] REIDEMEISTER, K.: Knotentheorie. Berlin 1932. — [5] REIDEMEISTER, K.: Homotopieringe und Linsenräume. Abh. Math. Sem. Hamburg **11**, 102—109 (1936). — [6] SCHUBERT, H.: Über eine numerische Knoteninvariante. Math. Z. **61**, 245—288 (1954). — [7] SCHUBERT, H.: Knoten und Vollringe. Acta Math. **90**, 131—286 (1953). — [8] SCHUBERT, H.: Die eindeutige Zerlegbarkeit eines Knotens in Primknoten. Sitzgsber. Akad. Wiss. Heidelberg, math.-nat. Kl., 3. Abh. **1949**. — [9] SEIFERT, H.: Schlingknoten. Math. Z. **52**, 62—80 (1949). — [10] TORRES, G., and R. H. FOX: Dual presentations of the group of a knot. Ann. Math. (2) **59**, 211—218 (1954).

[29]) Siehe REIDEMEISTER [4].

Heidelberg, Mathematisches Institut der Universität

(Eingegangen am 8. Dezember 1955)

SHAH, S. M., and S. K. SINGH
Math. Zeitschr. Bd. 65, S. 171—174 (1956)

On the derivative of a meromorphic function with maximum defect

By

S. M. SHAH and S. K. SINGH

Let $f(z)$ be a meromorphic function (not a constant) of finite order and $T(r, f)$ the characteristic function of $f(z)$. It is known that [1, p. 104]

$$\limsup_{r\to\infty} \frac{T(r, f')}{T(r, f)} \leq 2 - \mu(\infty) - \delta(\infty). \tag{1}$$

If $\delta(a_1) = \delta(a_2) = 1$ where $|a_i| < \infty$, $i = 1, 2$, then [2] $T(r, f') \sim 2T(r, f)$ as $r \to \infty$. We prove here

THEOREM: *If* $|a_i| < \infty$, $0 < |b_i| < \infty$, $i = 1, 2, 3, \ldots$ *then*

$$\Delta(f', 0) \liminf_{r\to\infty} \frac{T(r, f')}{T(r, f)} \geq \sum_1^\infty \delta(a_i) \tag{2}$$

$$\Delta(f', 0)\{2 - \mu(\infty) - \delta(\infty)\} \geq \{1 + \Delta(f', 0) - \delta(f', 0)\} \sum_1^\infty \delta(a_i), \tag{3}$$

where

$$\Delta(f', 0) = 1 - \liminf_{r\to\infty} \frac{N\left(r, \frac{1}{f'}\right)}{T(r, f')}.$$

COROLLARY: *If* $\sum_1^\infty \delta(a_i) = 2$,

then

$$T(r, f') \sim 2T(r, f), \qquad \mu(\infty) = 0,$$

$$\delta(f', 0) = 1, \qquad \delta(f', \infty) = 0$$

and

$$2\sum_1^\infty \delta(f', b_i) \leq 1.$$

PROOF. Consider first the case when $f(z)$ is a rational function:

$$f(z) = \frac{a z^n + \cdots}{A z^m + \cdots}.$$

If $m < n$, then it is easily seen that

$$\sum_1^\infty \delta(a_i) = 0$$

and (2) and (3) obviously hold.

If $m > n$, then

$$T(r, f) \sim m \log r,$$

$$T(r, f') \sim 2m \log r,$$

$$\delta(\infty) = \mu(\infty) = 0,$$

$$\sum \delta(a_i) = \delta(0) = 1 - \frac{n}{m},$$

$$\Delta(f', 0) = \frac{m - n + 1}{2m}$$

and (2) and (3) hold.

The case $m = n$ can be dealt with in a similar way.

Let now $f(z)$ be a meromorphic function (not a rational function). It is known that there can be at most a countably infinite number of numbers α such that $\delta(\alpha) > 0$.

Let $E = (A_i)_1^\infty$, $F = (B_i)_1^\infty$ be two sets of numbers such that $|A_i| < \infty$, $0 < |B_i| < \infty$ for $i = 1, 2, 3, \ldots$ and such that if[1]) $d \neq \infty$, $d \in CE$ then $\delta(d) = \delta(f, d) = 0$ and such that if $D \neq 0$, $\neq \infty$, $D \in CF$, then $\delta(f', D) = 0$; $\delta(f, A_i)$ and $\delta(f', B_i)$ may be positive or zero.

Let $\sum_1^\infty \delta(A_i) = x$. Given $\varepsilon > 0$ we can choose $(A_i)_1^p$ and $(B_i)_1^q$ such that

$$\sum_{p+1}^{\infty} \delta(A_i) < \varepsilon, \quad \sum_{q+1}^{\infty} \delta(f', B_i) < \varepsilon.$$

Then [3, p. 298]

$$p < \frac{T(r, f')}{T(r, f)} + \sum_1^p \frac{N(r, A_i)}{T(r, f)} - \frac{N\left(r, \frac{1}{f'}\right)}{T(r, f)} + o(1), \tag{4}$$

$$p + \liminf_{r \to \infty} \frac{N\left(r, \frac{1}{f'}\right)}{T(r, f')} \frac{T(r, f')}{T(r, f)} \leq \liminf_{r \to \infty} \frac{T(r, f')}{T(r, f)} + \sum_1^p \limsup_{r \to \infty} \frac{N(r, A_i)}{T(r, f)}.$$

Set $\beta = \liminf_{r \to \infty} \frac{T(r, f')}{T(r, f)}$; we have

$$p + \beta \{1 - \Delta(f', 0)\} \leq \beta + \sum_1^p \{1 - \delta(A_i)\},$$

$$\sum_1^p \delta(A_i) \leq \beta \Delta(f', 0),$$

$$\sum_1^\infty \delta(a_i) \leq x \leq \beta \Delta(f', 0) + \varepsilon$$

and (2) follows. Further we have from (4)

$$p + \limsup_{r \to \infty} \frac{N\left(r, \frac{1}{f'}\right)}{T(r, f')} \liminf_{r \to \infty} \frac{T(r, f')}{T(r, f)} \leq \limsup_{r \to \infty} \frac{T(r, f')}{T(r, f)} + \sum_1^p \limsup_{r \to \infty} \frac{N(r, A_i)}{T(r, f)}.$$

[1]) By CE we denote the complement of set E.

Hence from (2)

$$\Delta(f',0)\,p+\{1-\delta(f',0)\}\sum_1^\infty\delta(A_i)\leq\Delta(f',0)\{2-\mu(\infty)-\delta(\infty)\}+\Delta(f',0)\{p-\sum_1^p\delta(A_i)\},$$

$$\{1-\delta(f',0)\}\sum_1^\infty\delta(A_i)\leq\Delta(f',0)\{2-\mu(\infty)-\delta(\infty)\}-\Delta(f',0)\sum_1^\infty\delta(A_i)+\varepsilon.$$

Hence

$$\{1+\Delta(f',0)-\delta(f',0)\}\sum_1^\infty\delta(a_i)\leq\{1+\Delta(f',0)-\delta(f',0)\}\sum_1^\infty\delta(A_i)\leq\Delta(f',0)\{2-\mu(\infty)-\delta(\infty)\}+\varepsilon$$

and (3) follows.

COROLLARY. Since $\Delta(f',0)\leq 1$ we have

$$\liminf_{r\to\infty}\frac{T(r,f')}{T(r,f)}\geq\sum_1^\infty\delta(a_i)=2$$

and so from (1) we get

$$T(r,f')\sim 2\,T(r,f)\quad\text{and}\quad\mu(\infty)=0.$$

Further we have

$$(p-1)\,T(r,f)<\sum_1^p\overline{N}(r,a_i)+\overline{N}(r,f)+O(\log r).$$

Hence

$$(p-1)\leq\limsup_{r\to\infty}\sum_1^p\frac{\overline{N}(r,a_i)}{T(r,f)}+\liminf_{r\to\infty}\frac{\overline{N}(r,f)}{T(r,f)}$$

$$\leq p-\sum_1^p\delta(a_i)+\liminf_{r\to\infty}\frac{\overline{N}(r,f)}{T(r,f)}$$

$$\leq p-2+\varepsilon+\liminf_{r\to\infty}\frac{\overline{N}(r,f)}{T(r,f)}.$$

Hence

$$\liminf_{r\to\infty}\frac{\overline{N}(r,f)}{T(r,f)}=1.$$

Now

$$N(r,f')=N(r,f)+\overline{N}(r,f)\sim 2\,T(r,f),\qquad T(r,f')\sim 2\,T(r,f),$$

and so

$$\delta(f',\infty)=1-\limsup_{r\to\infty}\frac{N(r,f')}{T(r,f')}=0.$$

From (2) we have $\Delta(f',0)=1$ and from (3)

$$2\delta(f',0)\geq 2\quad\text{and so}\quad\delta(f',0)=1.$$

Further [3, p. 303]

$$p\,q < \frac{\bar{N}(r,f)}{T(r,f)} + q\sum_{1}^{p}\frac{N(r,A_i)}{T(r,f)} + \sum_{1}^{q}\frac{N\left(r,\frac{1}{f'-B_i}\right)}{T(r,f)} -$$

$$-\left\{(q-1)\frac{N\left(r,\frac{1}{f'}\right)}{T(r,f)} + \frac{N\left(r,\frac{1}{f''}\right)}{T(r,f)}\right\} + o(1).$$

And so

$$p\,q \leq 1 + q\sum_{1}^{p}\limsup_{r\to\infty}\frac{N(r,A_i)}{T(r,f)} + 2\sum_{1}^{q}\limsup_{r\to\infty}\frac{N\left(r,\frac{1}{f'-B_i}\right)}{T(r,f')}$$

$$\leq 1 + q\left\{p - \sum_{1}^{p}\delta(A_i)\right\} + 2\left\{q - \sum_{1}^{q}\delta(f',B_i)\right\}$$

$$\leq 1 + p\,q - q(2-\varepsilon) + 2q - 2\sum_{1}^{q}\delta(f',B_i).$$

And so

$$2\sum_{1}^{\infty}\delta(f',B_i) \leq 1 + (q+2)\,\varepsilon$$

which gives

$$2\sum_{1}^{\infty}\delta(f',b_i) \leq 2\sum_{1}^{\infty}\delta(f',B_i) \leq 1.$$

References

[1] Nevanlinna, R.: Le theorème de Picard-Borel et la theorie des fonctions méromorphes. Paris 1929. — [2] Shah, S. M., and S. K. Singh: Borel's theorem on a-points and exceptional values of Entire and Meromorphic Functions. Math. Z. **59**, 88—93 (1953). — [3] Milloux, H.: Les dérivées des fonctions méromorphes et la théorie des défauts. Ann. Ecole Normale Superieure **64**, 289—316 (1947).

Department of Mathematics, Muslim University, Aligarh (India)

D. S. College, Aligarh

(Eingegangen am 2. Dezember 1955)

GRAUERT, H., u. R. REMMERT
Math. Zeitschr. Bd. 65, S. 175—194 (1956)

Plurisubharmonische Funktionen in komplexen Räumen

Von

HANS GRAUERT und REINHOLD REMMERT

Die von K. OKA [*10*] und P. LELONG [*9*] eingeführten plurisubharmonischen Funktionen haben sich als ein unentbehrliches Hilfsmittel in der Theorie der Holomorphiegebiete erwiesen. K. OKA konnte zeigen, daß in jedem unverzweigten, pseudokonvexen RIEMANNschen Gebiet G über dem C^n eine holomorphe Funktion existiert, deren analytisches Gebilde mit G übereinstimmt. Er benutzt in seinem Beweis wesentlich die Eigenschaften der Funktion $-\ln \delta(x)$, die in unverzweigten, pseudokonvexen Gebieten plurisubharmonisch ist. Dabei bezeichnet $\delta(x)$ die euklidische Randdistanzfunktion in G. Es waren also die Untersuchungen über plurisubharmonische Funktionen, die nach langen vergeblichen Versuchen schließlich zu einem Nachweis der Richtigkeit der LEVIschen Vermutung führten.

Man hat gelegentlich vermutet [*8*], daß sich — analog dem Verhalten holomorpher Funktionen — jede in einem unverzweigten RIEMANNschen Gebiete G plurisubharmonische Funktion zu einer in der Holomorphiehülle $\mathfrak{H}(G)$ plurisubharmonischen Funktion fortsetzen läßt. Da $-\ln \delta(x)$ bei Annäherung von x an den Rand von G gegen $+\infty$ strebt und daher sicher nicht fortsetzbar ist, wäre dadurch schon eine Lösung des LEVIschen Problems gegeben worden. Inzwischen wurde aber diese Vermutung widerlegt [*2*]. Jedoch blieb das Problem offen, ob nicht für plurisubharmonische Funktionen ein Analogon zu den bekannten RIEMANNschen Sätzen über hebbare Singularitäten holomorpher Funktionen gilt, was offenbar eine schwächere Aussage ist. Die Sätze von RIEMANN drücken nämlich aus, daß man jede holomorphe Funktion f, die in einer Umgebung einer analytischen Menge $A \neq G$ eines Gebietes G des C^n holomorph ist, in A hinein holomorph fortsetzen kann. Dabei muß man, falls A von der maximalen Dimension $n-1$ ist, noch zusätzlich die Voraussetzung machen, daß $|f|$ in der Nähe jedes Punktes von A beschränkt ist.

Wir werden zeigen[1]):

I. *Ist G ein Gebiet im C^n und $A \neq G$ eine analytische Menge in G, so ist jede in $G-A$ plurisubharmonische Funktion $p(\mathfrak{z})$, die in der Nähe eines jeden Punktes $\mathfrak{z} \in A$ nach oben beschränkt ist, auf genau eine Weise zu einer im gesamten Gebiet G plurisubharmonischen Funktion $\overset{\circ}{p}(\mathfrak{z})$ fortsetzbar.*

[1]) Die Resultate der vorliegenden Arbeit wurden in einer Comptes-Rendus-Note angekündigt; vgl. [*5*].

II. *Ist A eine höchstens $(n-2)$-dimensionale analytische Menge in G, so existiert die plurisubharmonische Fortsetzung von $p(\mathfrak{z})$ in G stets, auch ohne die einschränkende Voraussetzung der Beschränktheit nach oben.*

Diese beiden Aussagen lassen sich sofort verallgemeinern, indem an Stelle analytischer Mengen beliebige dünne Mengen von G zugelassen werden. *Dünne Mengen* sind abgeschlossene Mengen in G, die lokal in analytischen Mengen enthalten sind.

Zur Charakterisierung der Pseudokonvexität abstrakter Gebilde, der sog. *komplexen Räume*, scheint es zweckmäßig, den Begriff der plurisubharmonischen Funktion auch hier zu kennen. In der vorliegenden Arbeit wird eine Definition für plurisubharmonische Funktionen in solchen komplexen Räumen gegeben; es wird bewiesen, daß dieselbe für Gebiete des Zahlenraumes mit der üblichen Definition übereinstimmt.

Die Fortsetzungssätze werden nun gleich für komplexe Räume formuliert und bewiesen; dabei macht die Existenz nichtuniformisierbarer Punkte in diesen Räumen besondere Schwierigkeiten. Es wird der Satz von F. HIRZEBRUCH [7] über die Auflösbarkeit der nichtuniformisierbaren Punkte in zweidimensionalen komplexen Räumen mittels Modifikationen herangezogen; ferner wird der folgende interessante Hilfssatz benutzt:

Ist $\mathfrak{R}$ eine zusammenhängende, endlich-blättrige, analytisch-verzweigte Überlagerung eines Polyzylinders Z im C^n, so ist für fast alle eindimensionalen analytischen Ebenen E die Beschränkung der Überlagerung auf $E \cap Z$ eine zusammenhängende RIEMANNsche *Fläche* (hinsichtlich der genauen Formulierung vgl. § 3, Hilfssatz 3).

Die Fortsetzungssätze gestatten Anwendungen auf die Theorie der Modifikationen. Ferner gewinnt man mit ihrer Hilfe eine Verallgemeinerung des OKAschen Satzes über die Äquivalenz der unverzweigten Holomorphiegebiete mit den pseudokonvexen Gebieten. Hierüber soll in einer weiteren Arbeit berichtet werden (vgl. auch [5]).

§ 1. Der Begriff der plurisubharmonischen Funktion Formulierung der Fortsetzungssätze

1. In Gebieten G des n-dimensionalen komplexen Zahlenraumes C^n kennt man den Begriff der plurisubharmonischen Funktion, den wir hier in leichter Abänderung der Definition von P. LELONG [9] wie folgt fassen:

Def. 1a. Eine Funktion $p(\mathfrak{z})$, $\mathfrak{z}=(z_1, \ldots, z_n)$, in einem Gebiete G des C^n heißt plurisubharmonisch in G, wenn sie den folgenden Bedingungen genügt:

α) Die Werte von $p(\mathfrak{z})$ sind reelle Zahlen oder $-\infty$.

β) $p(\mathfrak{z})$ ist in jedem Punkt $\mathfrak{z}_0 \in G$ halbstetig nach oben: $\overline{\lim_{\mathfrak{z} \to \mathfrak{z}_0}} p(\mathfrak{z}) \leq p(\mathfrak{z}_0)$.

γ) Ist E irgendeine eindimensionale analytische Ebene im C^n, die G schneidet, so ist die Beschränkung von $p(\mathfrak{z})$ auf $G \cap E$ eine subharmonische Funktion in $G \cap E$.

Dabei nennt man bekanntlich eine in einer offenen Menge B der komplexen Zahlenebene C definierte Funktion $s(z)$ subharmonisch in B, wenn folgendes gilt:

α') *Die Werte von* $s(z)$ *sind reelle Zahlen oder* $-\infty$.

β') $s(z)$ *ist in jedem Punkt* $z^{(0)} \in B$ *halbstetig nach oben:* $\overline{\lim}_{z \to z^{(0)}} s(z) \leq s(z^{(0)})$.

γ') *Ist* W *ein beliebiges Gebiet in* B *und* $h(z)$ *eine beliebige in* W *harmonische Funktion, so besitzt die Funktion* $h(z) + s(z)$, *sofern sie nicht in* W *konstant ist, kein Maximum in* W *(Maximumprinzip).*

Offensichtlich ist jede harmonische Funktion, da sie dem Maximumprinzip genügt, auch subharmonisch. Ferner ist eine Funktion in einem Gebiete der Zahlenebene genau dann dort plurisubharmonisch, wenn sie dort subharmonisch ist.

2. Wir notieren einige bekannte Eigenschaften plurisubharmonischer Funktionen. Die mindestens zweimal stetig differenzierbaren plurisubharmonischen Funktionen können nach LELONG [9] auch wie folgt charakterisiert werden:

a) *Eine in einem Gebiete* G *des* C^n *mindestens zweimal stetig differenzierbare Funktion* $p(\mathfrak{z})$ *ist genau dann plurisubharmonisch in* G, *wenn die quadratische Form* $\sum_{i,j} \frac{\partial^2 p}{\partial z_i \partial \bar{z}_j} dz_i d\bar{z}_j$ *überall in* G *positiv semi-definit ist.*

Beliebige plurisubharmonische Funktionen lassen sich nach OKA [10] stets durch differenzierbare plurisubharmonische Funktionen approximieren in folgendem Sinne:

b) *Ist* $p(\mathfrak{z})$ *irgendeine plurisubharmonische Funktion in einem Gebiete* $G \subset C^n$, *so gibt es zu jedem relativ-kompakt in* G *liegenden Gebiet* G' *eine absteigende Folge von in* G' *k-mal stetig differenzierbaren plurisubharmonischen Funktionen* $p_\nu(\mathfrak{z})$, *die in* G' *gegen* $p(\mathfrak{z})$ *konvergiert* ($k \geq 1$, *beliebig*).

Eine Aussage über den Limes einer Folge plurisubharmonischer Funktionen macht (vgl. [8])

c) *Die Grenzfunktion einer absteigenden Folge von plurisubharmonischen Funktionen in einem Gebiete* G *des* C^n *ist stets eine in* G *plurisubharmonische Funktion.*

Ist $P = \{p_\iota(\mathfrak{z}),\ \iota \in I\}$, I eine Indexmenge, eine Familie von in einem Gebiete G des C^n reellwertigen Funktionen ($-\infty$ sei als Wert zugelassen), die im Innern von G gleichmäßig nach oben beschränkt sind, so verstehen wir unter der *oberen Einhüllenden* $\hat{P}(\mathfrak{z})$ von P die kleinste nach oben halbstetige Funktion, für die gilt: $\hat{P}(\mathfrak{z}) \geq p_\iota(\mathfrak{z})$ für alle $\mathfrak{z} \in G$ und alle $\iota \in I$.

Die obere Einhüllende existiert stets. Es gilt [8]:

d) *Die obere Einhüllende* $\hat{P}(\mathfrak{z})$ *einer Familie* $P = \{p_\iota(\mathfrak{z})\}$ *von in einem Gebiete* G *des* C^n *plurisubharmonischen, im Innern von* G *gleichmäßig nach oben beschränkten Funktionen* $p_\iota(\mathfrak{z})$ *ist eine plurisubharmonische Funktion in* G.

3. In diesem Abschnitt beweisen wir zwei Sätze über plurisubharmonische Funktionen, die wir in dieser Arbeit benötigen.

Die Bedingung γ) in Definition 1a läßt sich wie folgt abschwächen:

Satz 1. Eine in einem Gebiete G des C^n definierte Funktion $p(\mathfrak{z})$, die den Bedingungen α) und β) genügt, ist bereits dann plurisubharmonisch in G, wenn folgendes gilt:

Es gibt im Raum der eindimensionalen analytischen Ebenen durch G eine dichte Menge M, derart, daß die Beschränkung von $p(\mathfrak{z})$ auf jede Ebene $E \in M$ eine subharmonische Funktion in $G \cap E$ ist.

Beweis: Es sei $p(\mathfrak{z})$ eine beliebige Funktion in G, die den Voraussetzungen des Satzes genügt; es sei

$$E_0 : \{z_\nu = a_\nu^{(0)} t + b_\nu^{(0)}, \qquad \nu = 1, \ldots, n\}$$

irgendeine eindimensionale analytische Ebene, die G schneidet. Wir müssen zeigen, daß $p(a_1^{(0)} t + b_1^{(0)}, \ldots, a_n^{(0)} t + b_n^{(0)})$ subharmonisch in $G \cap E_0$ ist. Mit M_ε, $\varepsilon > 0$, werde die Menge aller Ebenen $E_\iota : \{z_\nu = a_\nu^{(\iota)} t + b_\nu^{(\iota)}, \ \nu = 1, \ldots, n\}$ bezeichnet, die zu M gehören und für die gilt: $|a_\nu^{(\iota)} - a_\nu^{(0)}| < \varepsilon$, $|b_\nu^{(\iota)} - b_\nu^{(0)}| < \varepsilon$, $\nu = 1, \ldots, n$. Da M dicht im Raum aller Ebenen liegt, ist M_ε für kein $\varepsilon > 0$ leer.

Wir bilden zu jedem ε die zugehörigen Funktionsmengen

$$P_\varepsilon = \{p_\iota(t) = p(a_1^{(\iota)} t + b_1^{(\iota)}, \ldots, a_n^{(\iota)} t + b_n^{(\iota)})\}.$$

Die auftretenden Funktionen sind in ihren Definitionsbereichen sämtlich nach Voraussetzung subharmonisch und offensichtlich im Innern derselben gleichmäßig nach oben beschränkt. Nach 2.d) ist daher die obere Einhüllende $\hat{P}_\varepsilon(t)$ einer jeden Menge P_ε eine in ihrem Definitionsbereich subharmonische Funktion. Wir betrachten insbesondere die Folge $\hat{P}_{1/x}(t)$. Dieselbe ist monoton fallend; daher ist nach 2.c) die Grenzfunktion $\lim\limits_{x \to \infty} \hat{P}_{1/x}(t)$ ebenfalls subharmonisch. Da offensichtlich gilt: $\lim\limits_{x \to \infty} \hat{P}_{1/x}(t) = p(a_1^{(0)} \cdot t + b_1^{(0)}, \ldots, a_n^{(0)} \cdot t + b_n^{(0)})$, ist somit Satz 1 bewiesen.

Um den im folgenden Abschnitt einzuführenden Begriff der plurisubharmonischen Funktion in einem beliebigen komplexen Raum zu motivieren, beweisen wir noch:

Satz 2. Ist p irgendeine in einem Gebiete G des C^n plurisubharmonische Funktion, und ist τ irgendeine holomorphe Abbildung eines Gebietes $W \subset C$ in G, so ist $p \circ \tau$ stets eine subharmonische Funktion in W.

Wir führen den Beweis zunächst für mindestens zweimal stetig differenzierbare plurisubharmonische Funktionen p. Für diese ist nach 2.a) die Form $\sum\limits_{i,j} \frac{\partial^2 p}{\partial z_i \partial \bar{z}_j} dz_i d\bar{z}_j$ positiv semi-definit. Wird τ durch die in W holomorphen Funktionen

$$z_\nu = f_\nu(z), \qquad \nu = 1, \ldots, n, \ z \in W$$

beschrieben, so ist $p \circ \tau$ in W ebenfalls mindestens zweimal stetig differenzierbar, und $\frac{d^2 p \circ \tau(z)}{dz\, d\bar{z}} = \sum_{i,j} \frac{\partial^2 p}{\partial z_i \partial \bar{z}_j} \frac{df_i}{dz} \frac{\overline{df_j}}{dz}$ ist in W nicht negativ. Das besagt aber, daß $p \circ \tau$ subharmonisch in W ist.

Es sei nun p irgendeine in G plurisubharmonische Funktion. Es sei G_ν eine Folge von relativ-kompakt in G liegenden Gebieten, die G ausschöpft. Wir können nach 2.b) in jedem G_ν eine absteigende, gegen p konvergierende Folge von in G_ν zweimal stetig differenzierbaren, plurisubharmonischen Funktionen $p_\varkappa$ finden. Aus dem bereits Bewiesenen folgt, daß $p_\varkappa \circ \tau$ stets subharmonisch in $W_\nu = \tau^{-1}(\tau(W) \cap G_\nu)$ ist; dabei ist W_ν eine offene Menge in W. Nach 2.c) ist daher $p \circ \tau$ selbst subharmonisch in W_ν. Da die W_ν die Menge W ausschöpfen, folgt unmittelbar aus der Definition der subharmonischen Funktionen, daß $p \circ \tau$ auch in W selbst subharmonisch ist. — Satz 2 ist bewiesen.

4. Wir führen nun den Begriff der plurisubharmonischen Funktion in beliebigen komplexen Räumen ein. Ausgehend vom Begriff der endlichblättrigen analytisch-verzweigten Überlagerung denken wir uns den Begriff des komplexen Raumes wie in [4], [6] definiert. Ebenso seien analytische Mengen sowie holomorphe Abbildungen von komplexen Räumen ineinander wie in [4], [6] erklärt.

Def. 1. Eine Funktion $p(x)$ in einem komplexen Raum X heißt plurisubharmonisch in X, wenn sie den folgenden Bedingungen genügt:

α) Die Werte von $p(x)$ sind reelle Zahlen oder $-\infty$.

β) $p(x)$ ist in jedem Punkt $x_0 \in X$ halbstetig nach oben: $\overline{\lim_{x \to x_0}} p(x) \leq p(x_0)$.

γ) Ist τ irgendeine holomorphe Abbildung eines Gebietes W der Zahlenebene in den Raum X, so ist $p \circ \tau$ eine in W subharmonische Funktion.

Offensichtlich ist im Falle, daß X ein Gebiet G des C^n ist, jede nach Def. 1 plurisubharmonische Funktion in G auch nach Def. 1a plurisubharmonisch in G. Aus Satz 2 folgt sofort, daß auch die Umkehrung gilt:

Die Definitionen 1a und 1 sind für Gebiete im C^n äquivalent.

5. Für plurisubharmonische Funktionen in beliebigen komplexen Räumen gelten mutatis mutandis ebenfalls die in 2. gemachten Aussagen a) bis d). Wir notieren einige weitere Eigenschaften. Nützlich für viele Überlegungen ist:

e) *Eine Funktion p in einem komplexen Raum X ist genau dann plurisubharmonisch in X, wenn sie in jedem Punkte $x \in X$ plurisubharmonisch ist.*

Dabei nennen wir p plurisubharmonisch im Punkte x, wenn es eine Umgebung U von x gibt, so daß p im komplexen Raum U plurisubharmonisch ist.

Weiter gilt, wie für subharmonische Funktionen:

f) *Sind p und q plurisubharmonische Funktionen in einem komplexen Raum X, so ist auch die Summe $p+q$ eine plurisubharmonische Funktion in X.*

Dagegen ist die Differenz $p-q$ im allgemeinen nicht mehr plurisubharmonisch in X. Die Gesamtheit aller in einem komplexen Raum X plurisubharmonischen Funktionen bildet daher bezüglich der Addition eine Halbgruppe; sie sei gegebenenfalls mit $\Pi(X)$ bezeichnet.

Der Begriff der plurisubharmonischen Funktion ist invariant gegenüber holomorphen Abbildungen in folgendem Sinne:

g) *Ist p eine plurisubharmonische Funktion in einem komplexen Raum X und ist $\tau\colon Y \to X$ eine holomorphe Abbildung eines komplexen Raumes Y in X, so ist $p \circ \tau$ eine plurisubharmonische Funktion in Y.*

Der Abbildung $\tau\colon Y \to X$ entspricht somit in natürlicher Weise ein *Homomorphismus* $\tau^*\colon \Pi(X) \to \Pi(Y)$. Ist τ eine Abbildung auf X, so ist τ^* ein *Isomorphismus in* $\Pi(Y)$. Ist τ außerdem noch umkehrbar, so ist τ^* ein *Isomorphismus auf* $\Pi(Y)$.

Aus dem für subharmonische Funktionen gültigen Maximumprinzip folgt:

h) *Maximumprinzip für plurisubharmonische Funktionen: Nimmt eine in einem (zusammenhängenden) komplexen Raum X plurisubharmonische Funktion ihr Maximum in X an, so ist sie konstant.*

Hieraus erhält man:

i) *Verschärfung von Bedingung β): Ist p eine plurisubharmonische Funktion in einem komplexen Raum X, so gilt in jedem Punkte $x_0 \in X$: $\overline{\lim}_{x \to x_0} p(x) = p(x_0)$.*

In der Tat! Gäbe es einen Punkt $x_0 \in X$, so daß gelten würde: $\overline{\lim}_{x \to x_0} p(x) < p(x_0)$, so gäbe es eine zusammenhängende Umgebung U von x_0, derart, daß die Beschränkung $\hat{p}$ von p auf U ihr Maximum in x_0 annähme. Da $\hat{p}$ in U plurisubharmonisch und wegen $\hat{p}(x_0) > \overline{\lim}_{x \to x_0} \hat{p}(x)$ nicht konstant ist, hat man einen Widerspruch zum Maximumprinzip.

6. Wir wollen uns in dieser Arbeit mit Fortsetzungsfragen für plurisubharmonische Funktionen beschäftigen. Wir definieren:

Def. 2. Es sei X ein komplexer Raum; es sei $N \neq X$ eine abgeschlossene Menge in X. Eine in $X - N$ plurisubharmonische Funktion p heißt plurisubharmonisch fortsetzbar in X, wenn es in X eine plurisubharmonische Funktion $\breve{p}$ gibt, die in $X - N$ mit p übereinstimmt.

Ist p in X plurisubharmonisch fortsetzbar, so heißen die Punkte von N *hebbare Singularitäten* von p. Die plurisubharmonische Fortsetzung einer plurisubharmonischen Funktion ist im allgemeinen nicht eindeutig.

Zur Formulierung der Hauptresultate benötigen wir den Begriff der dünnen Menge[2]).

Def. 3. Eine abgeschlossene Teilmenge D eines n-dimensionalen komplexen Raumes X heißt dünn von der Ordnung k $(1 \leq k \leq n)$, wenn es zu jedem Punkt $x \in D$ eine Umgebung $U(x)$ und eine in U analytische, höchstens $(n-k)$-dimensionale Menge M gibt, für die gilt: $D \cap U \subset M$.

Hier bedarf der Begriff der Dimension einer analytischen Menge einer Erläuterung. Im Zahlenraum C^n sowie in komplexen Mannigfaltigkeiten sind mehrere Definitionen geläufig und äquivalent [3], [11]. Hinsichtlich der

[2]) Sinngemäßer wäre vielleicht die Bezeichnung analytisch dünne Menge. Der bequemeren Schreibweise halber haben wir das Beiwort „analytisch" unterdrückt.

präzisen Definition dieses Begriffes für komplexe Räume sei auf [6] verwiesen; für die Belange der vorliegenden Arbeit genügt folgende Aussage, die man etwa als Definition auffassen mag:

Eine in einem komplexen Raum X analytische Menge M ist im Punkte $x \in M$ genau dann höchstens s-dimensional, wenn folgendes gilt: Eine Umgebung $U(x) \subset X$ läßt sich durch eine holomorphe Abbildung τ eineindeutig auf eine endlich-blättrige, analytisch-verzweigte Überlagerung (R, Φ, G)[3]) eines Gebietes G des C^n abbilden, derart, daß $\Phi \circ \tau(M \cap U)$ in G in einer analytischen Menge von der Dimension $k \leq s$ enthalten ist. — Es ist leicht einzusehen, daß diese Definition unabhängig von (R, Φ, G) und τ ist.

7. In den nachstehenden Paragraphen beweisen wir die folgenden beiden Hauptsätze über hebbare Singularitäten plurisubharmonischer Funktionen.

Satz 3 (Analogon zum ersten Riemannschen *Satz über hebbare Singularitäten). Es sei D eine dünne Menge der Ordnung 1 in einem komplexen Raum X; es sei p eine in $X - D$ plurisubharmonische Funktion. Zu jedem Punkt $x \in D$ gebe es eine Umgebung U, derart, daß p in $U - U \cap D$ nach oben beschränkt ist. Dann ist p eindeutig zu einer in ganz X plurisubharmonischen Funktion $\breve{p}$ fortsetzbar.*

Satz 4 (Analogon zum zweiten Riemannschen *Satz über hebbare Singularitäten). Ist D eine dünne Menge der Ordnung 2 in einem komplexen Raum X, so ist jede in $X - D$ plurisubharmonische Funktion p eindeutig zu einer in ganz X plurisubharmonischen Funktion $\breve{p}$ fortsetzbar.*

Da die nichtuniformisierbaren Punkte eines komplexen Raumes X eine dünne Menge der Ordnung 2 in X sind (vgl. [6]), ergibt sich aus Satz 4 direkt:

Ist die Funktion p in jedem uniformisierbaren Punkt eines komplexen Raumes X plurisubharmonisch, so existiert genau eine plurisubharmonische Fortsetzung von p in ganz X.

Es sei angemerkt, daß P. Lelong mit völlig anderen Methoden als die in dieser Arbeit benutzten eine wesentliche Verallgemeinerung des Satzes 3 für den Fall, daß X eine komplexe Mannigfaltigkeit ist, bewiesen hat[4]).

Die Sätze 3 und 4 lassen sich besonders elegant formulieren, wenn man ihnen eine algebraische Fassung gibt. Ist $N \neq X$ eine abgeschlossene Teilmenge eines komplexen Raumes X, so bezeichne i^*: $\Pi(X) \to \Pi(X - N)$ den der Injektion i: $X - N \to X$ in kanonischer Weise zugeordneten Homomorphismus der Halbgruppe $\Pi(X)$ der in X plurisubharmonischen Funktionen in die Halbgruppe $\Pi(X - N)$ der in $X - N$ plurisubharmonischen Funktionen. Mit $\Pi^*(X - N)$ werde diejenige Unterhalbgruppe von $\Pi(X - N)$ bezeichnet,

[3]) Für analytisch-verzweigte Überlagerungen hat sich die Tripelschreibweise (R, Φ, G) als zweckmäßig erwiesen: R bezeichnet den eigentlichen Überlagerungsraum; G das Gebiet im C^n, über dem der Raum R liegt; und Φ die natürliche Abbildung von R auf G. Hinsichtlich der näheren Beschreibung vgl. [4] und [6].

[4]) Die Arbeit von Lelong wird in den Math. Ann. erscheinen. Vgl. auch C. R. Acad. Paris **242**, 55—57 (1956).

deren Elemente in der Umgebung eines jeden Punktes von N nach oben beschränkt sind. Dann gilt:

Satz 3a: Ist D eine dünne Menge der Ordnung 1 in einem komplexen Raum X, so ist der Homomorphismus $i^*: \Pi(X) \to \Pi(X-D)$ ein Isomorphismus von $\Pi(X)$ auf $\Pi^*(X-D)$.

Satz 4 formuliert sich wie folgt:

Satz 4a: Ist D eine dünne Menge der Ordnung 2 in einem komplexen Raum X, so ist der Homomorphismus $i^*: \Pi(X) \to \Pi(X-D)$ ein Isomorphismus von $\Pi(X)$ auf $\Pi(X-D)$.

§ 2. Beweis der Fortsetzungssätze für komplexe Mannigfaltigkeiten

1. Grundlegend für den Beweis der Sätze 3 und 4 ist der folgende Fortsetzungssatz für subharmonische Funktionen, der zugleich den Satz 3 für den Fall der Dimension 1 enthält.

Satz 5. Es sei B ein Bereich der z-Ebene; es sei $z^{(0)}$ irgendein Punkt aus B. In $B-z^{(0)}$ sei eine subharmonische Funktion $s(z)$ gegeben, die in einer Umgebung von $z^{(0)}$ nach oben beschränkt ist. Dann ist

$$\breve{s}(z) = \begin{cases} s(z) & \text{für} \quad z \neq z^{(0)}, \\ \overline{\lim\limits_{z \to z^{(0)}}}\, s(z) & \text{für} \quad z = z^{(0)}, \end{cases}$$

die eindeutig bestimmte subharmonische Fortsetzung von $s(z)$ in B[5]*.*

Beweis. Die Funktion $\breve{s}(z)$ ist wegen der Beschränktheit von $s(z)$ nach oben in der Umgebung von $z^{(0)}$ sinnvoll definiert: es gilt stets $\breve{s}(z) < +\infty$. Da $\breve{s}(z)$ nach Definition halbstetig nach oben ist in jedem Punkt $z \in B$, ist also nur noch zu zeigen, daß auch die Bedingung γ') erfüllt ist.

Es werde angenommen, daß dies nicht der Fall ist. Es gibt dann ein Teilgebiet W von B und eine in W harmonische Funktion h, derart, daß die Funktion $\breve{s}+h$ in W nicht konstant ist und ihr Maximum m in einem inneren Punkt $\tilde{z}^{(0)} \in W$ annimmt. Offenbar enthält W dann den Punkt $z^{(0)}$, und es gilt $\tilde{z}^{(0)} = z^{(0)}$; weiter wird der Wert m von $\breve{s}+h$ in keinem von $z^{(0)}$ verschiedenen Punkt aus W angenommen. Wir legen nun um $z^{(0)}$ einen Kreis $K_0: \{|z-z^{(0)}| \leq d_0\}$, der in W enthalten ist. Für jeden Punkt $z \in \{|z-z^{(0)}| = d_0\}$ gilt dann: $\breve{s}(z)+h(z) < m$. Da $\breve{s}+h$ halbstetig nach oben in W ist, können wir ein $\varepsilon > 0$ so finden, daß für jeden Punkt von $|z-z^{(0)}| = d_0$ sogar gilt: $\breve{s}(z)+h(z) < m-\varepsilon$.

Wir bilden jetzt im Kreisring $K_d: \{d < |z-z^{(0)}| < d_0\}$, $0 < d < d_0$, die harmonische Funktion

$$h_d(z) = \frac{\varepsilon}{\ln \dfrac{d}{d_0}} \cdot \ln \left\{ \left(\frac{d}{d_0} \right)^m \cdot \frac{|z-z^{(0)}|}{d} \right\}.$$

[5]) Dieser Satz wurde im wesentlichen bereits von BRELOT [1] im Jahre 1934 bewiesen. Der hier wiedergegebene Beweis stimmt nicht mit dem von BRELOT mitgeteilten überein.

Offensichtlich ist:

$$h_d(z) = \begin{cases} m & \text{für } z \in \{|z - z^{(0)}| = d\} \\ m - \varepsilon & \text{für } z \in \{|z - z^{(0)}| = d_0\}. \end{cases}$$

Daher gilt für jeden Punkt $z \in K_d$: $\breve{s}(z) + h(z) \leq h_d(z)$.

Wie man leicht nachrechnet, konvergiert die Folge $h_d(z)$ für $d \to 0$ im Innern von $K_0 - z^{(0)}$ gleichmäßig gegen $m - \varepsilon$. Es gilt also in $K_0 - z^{(0)}$:

$$\breve{s}(z) + h(z) \leq m - \varepsilon.$$

Das ist jedoch nicht mit der Gleichung $\breve{s}(z^{(0)}) + h(z^{(0)}) = m$ verträglich. Mithin ist $\breve{s}(z)$ eine in B subharmonische Funktion. Wegen der verschärften Bedingung β) ist klar, daß $\breve{s}(z)$ die einzige subharmonische Fortsetzung von $s(z)$ in B ist.

2. Der Beweis des Satzes 3 für komplexe Räume beliebiger Dimension bereitet größere Schwierigkeiten, die durch die Existenz nichtuniformisierbarer Punkte in solchen Räumen bedingt sind. Wir schließen daher zunächst diesen Fall aus und beweisen den folgenden Spezialfall von Satz 3:

Satz 6. Es sei D eine dünne Menge der Ordnung 1 in einer komplexen Mannigfaltigkeit X und $p(x)$ eine in $X - D$ plurisubharmonische Funktion, die in einer Umgebung eines jeden Punktes von D nach oben beschränkt ist. Dann ist

$$\breve{p}(x') = \begin{cases} p(x') & \text{für } x' \notin D, \\ \overline{\lim}_{x \to x',\ x \in X - D} p(x) & \text{für } x' \in D, \end{cases}$$

die eindeutig bestimmte plurisubharmonische Fortsetzung von $p(x)$ in X.

Anmerkung. Die Redeweise, p ist in einer Umgebung eines jeden Punktes $x \in D$ nach oben beschränkt, soll heißen, daß es eine Umgebung U von x gibt, so daß p in $U - U \cap D$ nach oben beschränkt ist.

Beweis. Die Funktion $\breve{p}(x)$ ist wegen der Beschränktheit von $p(x)$ nach oben in der Umgebung eines jeden Punktes von D in ganz X sinnvoll definiert; es gilt stets: $\breve{p}(x) < +\infty$. Weiter ist $\breve{p}(x)$ in jedem Punkt $x \in X$ halbstetig nach oben. Gäbe es nämlich einen Punkt $x^* \in D$ und eine gegen x^* konvergierende Folge $x_\nu \in X$ mit $\overline{\lim}_{x_\nu \to x^*} \breve{p}(x_\nu) > \breve{p}(x^*)$, so gäbe es, wie aus der Definition von $\breve{p}(x)$ unmittelbar folgt, auch eine Folge $\tilde{x}_\nu \in X - D$ mit derselben Eigenschaft; das widerspricht jedoch der Definition von $\breve{p}$ in x^*.

Es bleibt noch die Bedingung γ) zu verifizieren. Da alle Punkte von X uniformisierbar sind, kann man zu jedem Punkt $x \in D$ eine Umgebung U finden, die sich eineindeutig und holomorph auf ein Gebiet G des C^n abbilden läßt. Bezeichnet τ eine solche Abbildung, so ist nach § 1.5.g) die Funktion $\breve{p}(x)$ sicher dann plurisubharmonisch in X, wenn $\breve{p} \circ \tau^{-1} = (p \breve{\circ} \tau^{-1})$ plurisubharmonisch in G ist. Wir brauchen also Satz 6 nur für Gebiete G des C^n zu beweisen. Da wir wegen § 1.5.e) das Gebiet G beliebig klein wählen dürfen,

können wir noch annehmen, daß D im Nullstellengebilde F einer in G holomorphen, nicht identisch verschwindenden Funktion f enthalten ist. Es werde nun Satz 6 für diesen Fall bewiesen.

Nach Satz 1 genügt es, im Raum der eindimensionalen analytischen Ebenen durch G eine dichte Menge M anzugeben, derart, daß die Beschränkung von $\breve{p}$ auf jede Ebene $E \in M$ subharmonisch ist. Wir wählen für M die offensichtlich dichte Menge aller derjenigen eindimensionalen analytischen Ebenen durch G, die die in G analytische Menge $F: \{f = 0\}$ punkthaft schneiden.

Ist nun $\tilde{E}$ irgendeine Ebene aus M und $\mathfrak{z}_0$ ein isolierter Schnittpunkt von $\tilde{E}$ mit D (man beachte: $D \subset F$), so ist die Beschränkung von $\breve{p}$ auf $\tilde{E}$ in der Umgebung von $\mathfrak{z}_0$ auf Grund von Satz 5 sicher dann subharmonisch, wenn die Gleichung gilt:

$$\breve{p}(\mathfrak{z}_0) = \overline{\lim_{\mathfrak{z} \to \mathfrak{z}_0,\ \mathfrak{z} \in \tilde{E} - \mathfrak{z}_0}} p(\mathfrak{z}).$$

Sicher gilt $\overline{\lim}_{\mathfrak{z} \to \mathfrak{z}_0,\ \mathfrak{z} \in \tilde{E} - \mathfrak{z}_0} p(\mathfrak{z}) \leq \breve{p}(\mathfrak{z}_0)$; wir nehmen zur Erzielung eines Widerspruches an, es würde gelten: $\overline{\lim}_{\mathfrak{z} \to \mathfrak{z}_0,\ \mathfrak{z} \in \tilde{E} - \mathfrak{z}_0} p(\mathfrak{z}) < m < \breve{p}(\mathfrak{z}_0)$. Es sei dann $\mathfrak{z}_\nu \in G - D$ eine gegen $\mathfrak{z}_0$ konvergierende Punktfolge mit $\lim_{\nu \to \infty} p(\mathfrak{z}_\nu) = \breve{p}(\mathfrak{z}_0)$, welche es laut Definition von $\breve{p}(\mathfrak{z}_0)$ gibt. Mit E_ν werde diejenige durch $\mathfrak{z}_\nu$ laufende eindimensionale analytische Ebene bezeichnet, die parallel zu $\tilde{E}$ ist. Die Ebenen E_ν schneiden die Menge D bei hinreichend groß gewähltem Index ν ebenfalls punkthaft.

Bezeichnen wir mit $p_\nu(\mathfrak{z})$ die Beschränkung von $p(\mathfrak{z})$ auf $E_\nu - E_\nu \cap D$, so läßt sich also jede Funktion $p_\nu(\mathfrak{z})$ bei genügend groß gewähltem ν auf Grund von Satz 5 zu einer in ganz E_ν subharmonischen Funktion $\breve{p}_\nu(\mathfrak{z})$ fortsetzen [es wird an dieser Stelle nicht behauptet, daß $\breve{p}_\nu(\mathfrak{z})$ mit der Beschränkung von $\breve{p}(\mathfrak{z})$ auf E_ν übereinstimmt!]. Sei nun $\tilde{K}$ ein Kreis auf $\tilde{E}$ um $\mathfrak{z}_0$; seien K_ν Kreise von gleichem Radius r um $\mathfrak{z}_\nu$ auf E_ν. Für hinreichend kleinen Radius r und genügend großes ν haben die Ränder $\partial \tilde{K}$, ∂K_ν der Kreise $\tilde{K}$, K_ν mit der Menge D keinen Punkt gemeinsam; man kann erreichen, daß gilt:

$$\breve{p}_\nu(\mathfrak{z}) = p(\mathfrak{z}) < m \quad \text{für} \quad \mathfrak{z} \in \partial K_\nu,\ \nu > \nu_0.$$

Bei genügend kleinem r ist nämlich wegen der Annahme $\overline{\lim}_{\mathfrak{z} \to \mathfrak{z}_0,\ \mathfrak{z} \in \tilde{E} - \mathfrak{z}_0} p(\mathfrak{z}) < m$ auch $p(\mathfrak{z}) < m$ für $\mathfrak{z} \in \partial \tilde{K}$. Weil ferner $p(\mathfrak{z})$ halbstetig nach oben ist, gilt das noch in einer ganzen Umgebung von ∂K, also auch auf ∂K_ν, $\nu > \nu_0$. Für hinreichend großes ν ist aber auch

$$\breve{p}_\nu(\mathfrak{z}_\nu) > m.$$

Da $\mathfrak{z}_\nu \in K_\nu$, würde also die in K_ν subharmonische Funktion $\breve{p}_\nu$ ihr Maximum in K_ν annehmen und nicht konstant sein. Das widerspricht jedoch dem Maximumprinzip. Folglich gilt doch: $\overline{\lim}_{\mathfrak{z} \to \mathfrak{z}_0,\ \mathfrak{z} \in \tilde{E} - \mathfrak{z}_0} p(\mathfrak{z}) = \breve{p}(\mathfrak{z}_0)$, d.h. $\breve{p}(\mathfrak{z})$ ist eine plurisubharmonische Fortsetzung von $p(\mathfrak{z})$ in X.

Es bleibt zu zeigen, daß $\breve{p}(\mathfrak{z})$ die einzige plurisubharmonische Fortsetzung von $p(\mathfrak{z})$ in X ist. Zu dem Zwecke sei $\breve{p}'(\mathfrak{z})$ irgendeine weitere plurisubharmonische Fortsetzung von $p(\mathfrak{z})$ in X. Dann gilt $\breve{p}'(\mathfrak{z}) = \breve{p}(\mathfrak{z}) = p(\mathfrak{z})$ für jeden Punkt $\mathfrak{z} \in X - D$. Ist nun $\mathfrak{z}_0 \in D$ irgendein Punkt, so sei E eine eindimensionale RIEMANNsche Fläche durch $\mathfrak{z}_0$, die D in der Nähe von $\mathfrak{z}_0$ nur in $\mathfrak{z}_0$ schneidet. Da die Beschränkungen von $\breve{p}'$ und $\breve{p}$ auf E subharmonisch sind, folgt aus der verschärften Bedingung β') in beiden Fällen:

$$\overline{\lim_{\mathfrak{z} \to \mathfrak{z}_0,\ \mathfrak{z} \in E - \mathfrak{z}_0}} p(\mathfrak{z}) = \begin{cases} \breve{p}(\mathfrak{z}_0), \\ \breve{p}'(\mathfrak{z}_0). \end{cases}$$

Daher stimmt $\breve{p}'$ in jedem Punkt von X mit $\breve{p}$ überein. — Satz 6 ist bewiesen.

§ 3. Beweis der Fortsetzungssätze für beliebige komplexe Räume

1. Um den Allgemeinfall des Satzes 3 zu beweisen, sind umfangreiche Vorbereitungen erforderlich. Wir benötigen zunächst einen Fortsetzungssatz für endlich-blättrige, analytisch-verzweigte Überlagerungen (zum Beweise vgl. [6]):

Fortsetzungssatz. Ist A eine höchstens $(n-1)$-dimensionale analytische Menge in einem Gebiete G des C^n, so ist jede endlich-blättrige, analytisch-verzweigte Überlagerung $(R, \Phi, G - A)$ von $G - A$ eindeutig zu einer endlich-blättrigen, analytisch-verzweigten Überlagerung $(\breve{R}, \breve{\Phi}, G)$ von G fortsetzbar.

Weiter werden wir einen grundlegenden Satz von F. HIRZEBRUCH aus der Theorie der Modifikationen entscheidend verwenden (zur Theorie der Modifikationen vgl. [4]).

Ein komplexer Raum $'X$ heißt eine (eigentliche) Modifikation eines komplexen Raumes X, wenn es eine eigentliche holomorphe Abbildung τ von $'X$ auf X und dünne Mengen $N \subset X$, $'N \subset {}'X$ der Ordnung 1 gibt, so daß τ den Raum $'X - {}'N$ umkehrbar eindeutig auf $X - N$ abbildet.

F. HIRZEBRUCH hat bewiesen [7]:

Satz. Zu jedem zweidimensionalen komplexen Raum X gibt es eine eigentliche Modifikation $'X$, derart, daß $'X$ eine komplexe Mannigfaltigkeit und $N \subset X$ die Menge der nichtuniformisierbaren Punkte von X ist.

Bekanntlich liegen bei zweidimensionalen komplexen Räumen die nichtuniformisierbaren Punkte isoliert. Offensichtlich ist die Menge ${}'N_x \subset {}'X$, die vermöge der Modifikationsabbildung τ auf einen nichtuniformisierbaren Punkt $x \in X$ abgebildet wird, eine analytische Menge in $'X$. Wir zeigen nun für $n = 2$:

Hilfssatz 1. Die Menge ${}'N_x$ ist stets zusammenhängend.

Beweis. Wäre die Behauptung falsch, so zerfiele die Menge ${}'N_x$ für einen gewissen nichtuniformisierbaren Punkt $x \in X$ in zwei in $'X$ analytische Mengen ${}'N_1$ und ${}'N_2$, die punktfremd sind. Da ${}'N_1$ und ${}'N_2$ abgeschlossen sind, gibt es auch punktfremde zusammenhängende Umgebungen ${}'U_1({}'N_1)$ und ${}'U_2({}'N_2)$,

die keinen weiteren Punkt von $'N$ enthalten. Es ist nun

$$\tau\{('U_1('N_1) - 'N_1) \cup ('U_2('N_2) - 'N_2)\} \cup x$$

eine zusammenhängende Umgebung $\dot{U}$ des Punktes x, wie man sofort einsieht, wenn man beachtet, daß τ eine eigentliche Abbildung ist. Nun ist auch $U - x$ zusammenhängend. Andererseits ist aber $U - x$ die Vereinigung der beiden punktfremden offenen Mengen $\tau('U_1('N_1) - 'N_1)$ und $\tau('U_2('N_2) - 'N_2)$. Widerspruch! Der Hilfssatz ist bewiesen.

2. Es sei nun $\mathfrak{R} = (R, \Phi, Z^2)$ eine endlich-blättrige, analytisch-verzweigte Überlagerung eines zweidimensionalen Polyzylinders Z^2: $\{|z_1 - z_1^{(0)}| < d, |z_2 - z_2^{(0)}| < d\}$; über dem Punkt $\mathfrak{z}_0 = (z_1^{(0)}, z_2^{(0)}) \in Z^2$ liege nur ein Punkt $r_0 \in R$. Die Verzweigungspunkte von $\mathfrak{R}$ seien sämtlich über der in Z^2 analytischen eindimensionalen Menge V gelegen; die Ebene E: $\{z_2 = z_2^{(0)}\}$ schneide V nur im Punkte $\mathfrak{z}_0$. Wir betrachten die analytisch-verzweigte Überlagerung $(\Phi^{-1}(E \cap Z^2 - \mathfrak{z}_0), \Phi, E \cap Z^2 - \mathfrak{z}_0)$. Dieselbe zerfällt in endlich viele zusammenhängende Komponenten $\tilde{\mathfrak{R}}_\varkappa = (\tilde{R}_\varkappa, \tilde{\Phi}_\varkappa, E \cap Z^2 - \mathfrak{z}_0)$, $\varkappa = 1, \ldots, k$; der Punkt r_0 ist Häufungspunkt jeder Menge $\tilde{R}_\varkappa$. Wir behaupten:

Hilfssatz 2. Ist $p(r)$ eine auf $R - r_0$ plurisubharmonische Funktion, die in der Umgebung des Punktes r_0 nach oben beschränkt ist, so gilt:

$$\overline{\lim_{r \to r_0,\ r \in \tilde{R}_\varkappa}} p(r) = \overline{\lim_{r \to r_0,\ r \in R - r_0}} p(r) = m \quad \text{für jedes} \quad \varkappa = 1, \ldots, k.$$

Beweis. Sicher existieren sämtliche angeschriebenen limites superiores. Um zu zeigen, daß sie alle übereinstimmen, modifizieren wir nach F. HIRZEBRUCH den zweidimensionalen komplexen Raum R zu einer komplexen Mannigfaltigkeit $'R$. Die Funktion $'p = p \circ \tau$, wo τ die zugehörige Modifikationsabbildung bezeichnet, ist nun in $'R - 'N_{r_0}$, wo $'N_{r_0} = \tau^{-1}(r_0)$, plurisubharmonisch und bleibt in der Umgebung von $'N_{r_0}$ nach oben beschränkt, da p in der Umgebung von r_0 nach oben beschränkt ist. Da $'N_{r_0}$ insbesondere eine dünne Menge der Ordnung 1 in der komplexen Mannigfaltigkeit $'R$ ist, so läßt sich $'p$ nach Satz 6 zu einer in ganz $'R$ plurisubharmonischen Funktion $'\breve{p}$ fortsetzen. Die kompakte Menge $'N_{r_0}$ zerfällt in $'R$ in endlich viele algebraische RIEMANNsche Flächen $'A_1, \ldots, 'A_l$. Aus dem Maximumprinzip für subharmonische Funktionen folgt, daß $'\breve{p}$ auf allen kompakten RIEMANNschen Flächen $'A_1, \ldots, 'A_l$ konstant ist. Nach Hilfssatz 1 ist $'N_{r_0} = \bigcup_{\lambda=1}^{l} 'A_\lambda$ zusammenhängend; daher ist der Wert von $'\breve{p}$ auf allen $'A_\lambda$ derselbe und mithin $'\breve{p}$ auf $'N_{r_0}$ selbst konstant. Die in $'R - 'N_{r_0}$ analytische Menge $\tau^{-1}(\tilde{R}_\varkappa)$ ist für jedes $\varkappa$ zu einer in $'R$ irreduziblen RIEMANNschen Fläche $'\tilde{R}_\varkappa$ fortsetzbar[6]; dieselbe hat dann mit $'N_{r_0}$ genau einen Punkt $'r_\varkappa$ gemeinsam, der eventuell der einzige nicht gewöhnliche Punkt von $'\tilde{R}_\varkappa$ ist. Da die Beschränkung von

[6]) Die Fortsetzbarkeit jeder Menge $\tau^{-1}(\tilde{R}_\varkappa)$ zu einer in $'R$ irreduziblen RIEMANNschen Fläche folgert man am einfachsten aus einem allgemeinen Fortsetzungssatz von analytischen Mengen bei eigentlichen Modifikationen, vgl. [4].

$'\breve{p}$ auf $'\tilde{R}_\varkappa$ eine subharmonische Funktion ist, folgt aus § 1.5.i), daß gilt:

$$\overline{\lim_{r \to r_0,\ r \in \tilde{R}_\varkappa}} p(r) = \lim_{'r \to 'r_\varkappa,\ 'r \in 'R_\varkappa - 'r_\varkappa} p \circ \tau('r) = '\breve{p}('r_\varkappa), \quad \varkappa = 1, \ldots, k.$$

Da $'r_\varkappa \in 'N_{r_0}$ und $'\breve{p}$ auf $'N_{r_0}$ konstant ist, folgt:

$$'\breve{p}('r_\varkappa) = \overline{\lim_{'r \to 'N_{r_0},\ 'r \subset 'R - 'N_{r_0}}} 'p('r) = m.$$

Der hier rechts stehende Ausdruck stimmt aber wegen $'p = p \circ \tau$ mit $\overline{\lim_{r \to r_0,\ r \subset R - r_0}} p(r)$ überein. Somit ist Hilfssatz 2 bewiesen.

3. In diesem Abschnitt betrachten wir endlich-blättrige, analytisch-verzweigte Überlagerungen $\mathfrak{R} = (R, \Phi, Z)$ eines n-dimensionalen Polyzylinders Z: $\{|z_1 - z_1^{(0)}| < d, \ldots, |z_n - z_n^{(0)}| < d\}$, $n \geq 2$. Analog wie in Abschnitt 2 werde angenommen, daß über $\mathfrak{z}_0 = (z_1^{(0)}, \ldots, z_n^{(0)})$ nur ein Punkt $r_0 \in R$ liegt; weiter sei V eine in Z analytische, rein $(n-1)$-dimensionale Menge, über der sämtliche Verzweigungspunkte von $\mathfrak{R}$ liegen. Wir setzen voraus, daß V enthalten ist im Polyzylinder Z': $\left\{|z_1 - z_1^{(0)}| < \frac{d}{2}\right\} \times \hat{Z}$, wobei wir setzen:

$$\hat{Z}: \{|z_2 - z_2^{(0)}| < d, \ldots, |z_n - z_n^{(0)}| < d\}.$$

Alsdann ist V das Nullstellengebilde eines Pseudopolynoms

$$\omega = z_1^s + A_1(z_2, \ldots, z_n) z_1^{s-1} + \cdots + A_s(z_2, \ldots, z_n)$$

mit in $\hat{Z}$ holomorphen Koeffizienten $A_\sigma(z_2, \ldots, z_n)$; man kann annehmen, daß ω keine mehrfachen Faktoren besitzt. Es bezeichne $\hat{V}$ die Diskriminantenmenge von ω im $(n-1)$-dimensionalen Polyzylinder $\hat{Z}$; $\hat{V}$ ist eine analytische, rein $(n-2)$-dimensionale oder leere Menge in $\hat{Z}$.

Wir behaupten:

Hilfssatz 3. Ist E^1: $\{z_2 = a_2, \ldots, z_n = a_n\}$ *irgendeine eindimensionale analytische Ebene in* Z *mit* $(a_2, \ldots, a_n) \in \hat{Z} - \hat{V}$, *so ist* $(\Phi^{-1}(E^1 \cap Z - V), \Phi, E^1 \cap Z - V)$ *eine zusammenhängende unverzweigte Überlagerung von* $E^1 \cap Z - V$.

Beweis. Sicher ist $E^1 \cap Z - V$ und damit auch $\tilde{R} = \Phi^{-1}(E^1 \cap Z - V)$ nicht leer. Da $\tilde{R}$ keine Punkte über V besitzt, ist Φ in $\tilde{R}$ eine lokal-topologische Abbildung; daher ist $(\tilde{R}, \Phi, E^1 \cap Z - V)$ eine unverzweigte Überlagerung von $E^1 \cap Z - V$. Wäre nun der Raum $\tilde{R}$ nicht zusammenhängend, so zerfiele er in mindestens zwei zusammenhängende Komponenten; seien $\tilde{R}_0$ und $\tilde{R}_1$ solche. Man kann eine Kurve $\{q(t), 0 \leq t \leq 1\}$ in R finden, die einen Punkt aus $\tilde{R}_0$ mit einem Punkt aus $\tilde{R}_1$ verbindet, so daß folgendes gilt:

a) $\{q(t), 0 \leq t \leq 1\}$ verläuft in $R - \Phi^{-1}(V)$.

b) Bezeichnet $\{z_1(t), \ldots, z_n(t), 0 \leq t \leq 1\}$ die Kurve $\{\Phi \circ q(t), 0 \leq t \leq 1\}$, so verläuft $\{\hat{q}(t), 0 \leq t \leq 1\} = \{z_2(t), \ldots, z_n(t), 0 \leq t \leq 1\}$ in $\hat{Z} - \hat{V}$.

Wir betrachten nun die Schar $\{E_t^1\}$ der eindimensionalen analytischen Ebenen E_t^1: $\{z_2 = z_2(t), \ldots, z_n = z_n(t)\}$, und bezeichnen mit T den topologischen

Raum der Paare $(\mathfrak{z}, t)$, wo $\mathfrak{z} \in E_t^1 \cap Z - V$ und $0 \leq t \leq 1$. Wir behaupten: *T ist homöomorph dem kartesischen Produkt eines k-fach punktierten Kreises mit dem Einheitsintervall.*

Da laut Voraussetzung die Kurve $\{\hat{\varphi}(t)\}$ keinen Punkt mit der Diskriminantenmenge $\hat{V}$ von V gemeinsam hat, schneidet jede Ebene E_t^1 die Menge V, die sich als eine Überlagerung von $\hat{Z}$ auffassen läßt, stets außerhalb der Verzweigungspunkte von V über $\hat{Z}$. Es gibt folglich endlich viele stetige Funktionen $s_1(t), \ldots, s_k(t)$, derart, daß $s_\varkappa(t) \neq s_\lambda(t)$ für $\varkappa \neq \lambda$ und $\bigcup_{\varkappa=1}^{k} s_\varkappa(t) = E_t^1 \cap V$. Daher besteht T aus allen Paaren $(\mathfrak{z}, t)$ mit $\mathfrak{z} \in E_t^1 \cap Z$, $\mathfrak{z} \neq s_\varkappa(t)$, $\varkappa = 1, \ldots, k$, $0 \leq t \leq 1$. Daraus ergibt sich aber direkt die Behauptung.

Neben dem Raum T betrachten wir weiter den Raum R^* der Paare (r, t), wo $r \in \Phi^{-1}(E_t^1 \cap Z - V)$ und $0 \leq t \leq 1$. Setzt man $\Phi^*((r, t)) = (\Phi(r), t)$, so ist offenbar (R^*, Φ^*, T) eine endlich-blättrige unbegrenzte Überlagerung von T im topologischen Sinne. Dieselbe ist unverzweigt, da Φ^* überall lokal-topologisch ist. Es sei R_0^* diejenige zusammenhängende Komponente von R^*, die die Kurve $\{\varphi^*(t), 0 \leq t \leq 1\} = \{(\varphi(t), t);\ 0 \leq t \leq 1\}$ enthält. Ferner werde gesetzt: $\tilde{R}_t^* = R_0^* \cap \Phi^{*-1}((E_t^1 \cap Z - V, t))$; wir behaupten, daß $\tilde{R}_t^*$ stets zusammenhängend ist. Da nämlich T homöomorph dem kartesischen Produkt eines k-fach punktierten Kreises mit dem Einheitsintervall ist, läßt sich in T jede geschlossene Kurve stetig in eine Kurve deformieren, die in einer der Ebenen $\left(E_t^1 - \bigcup_{\varkappa=1}^{k} s_\varkappa(t), t\right)$ liegt; dabei darf t zwischen 0 und 1 beliebig gewählt werden.

Aus dieser Eigenschaft von T ergibt sich aber sofort der Zusammenhang jeder Menge $\tilde{R}_t^*$, wenn man beachtet, daß (R_0^*, Φ^*, T) eine zusammenhängende, unbegrenzte, unverzweigte Überlagerung von T ist; denn man überlegt sofort, daß allgemein gilt:

Ist (E, Ψ, F) eine unverzweigte, unbegrenzte, zusammenhängende Überlagerung eines zusammenhängenden, lokal-zusammenhängenden topologischen Raumes F, und $F' \subset F$ eine zusammenhängende und lokal-zusammenhängende Menge in F, derart, daß jede in F verlaufende, geschlossene Kurve stetig deformierbar ist in eine geschlossene Kurve, die in F' verläuft, so ist $(\Psi^{-1}(F'), \Psi, F')$ zusammenhängend.

Der Raum $\tilde{R}_0^*$ enthält den Anfangspunkt und der Raum $\tilde{R}_1^*$ den Endpunkt der Kurve $\{\varphi^*(t)\}$. Bezeichnen wir mit π: $(r, t) \to r$ die natürliche, stetige Abbildung von R^* in R, so wird also mindestens ein Punkt aus $\tilde{R}_0^*$ vermöge π in $\tilde{R}_0$ und mindestens ein Punkt von $\tilde{R}_1^*$ in $\tilde{R}_1$ abgebildet. Da aber $\tilde{R}_0^*$ und $\tilde{R}_1^*$ nach dem Vorstehenden zusammenhängend sind und durch π beide in R abgebildet werden, so folgt, daß $\tilde{R}_0^*$ durch π ganz in $\tilde{R}_0$ und $\tilde{R}_1^*$ durch π ganz in $\tilde{R}_1$ abgebildet wird.

Wir betrachten nun die Überlagerung $(\Phi^{-1}(Z - \overline{Z}'), \Phi, Z - \overline{Z}')$ (der Polyzylinder Z' war definiert durch $\left\{|z_1 - z_1^{(0)}| < \frac{d}{2}\right\} \times \hat{Z}$); dieselbe ist unverzweigt

und unbegrenzt über $Z-\overline{Z}'$ und zerfällt in endlich viele zusammenhängende Komponenten $(B_\lambda, \Phi, Z-\overline{Z}')$, $\lambda=0,1,\ldots,l$. Da sich überdies jede in $Z-\overline{Z}'$ verlaufende geschlossene Kurve stetig deformieren läßt in eine geschlossene Kurve, die in einer beliebig vorgegebenen, $Z-\overline{Z}'$ schneidenden eindimensionalen analytischen Ebene liegt, folgt analog wie oben, daß das Φ-Urbild einer jeden in $Z-\overline{Z}'$ eindimensionalen analytischen Ebene jeweils eine zusammenhängende Menge in B_λ ist. Daher können $\widetilde{R}_0\cap\Phi^{-1}(Z-\overline{Z}')$ und $\widetilde{R}_1\cap\Phi^{-1}(Z-\overline{Z}')$ nicht beide Punkte in derselben Komponente B_λ besitzen; die Numerierung der B_λ kann folglich so gewählt werden, daß $\widetilde{R}_0\cap\Phi^{-1}(Z-\overline{Z}')$ in der Menge $C=\bigcup_{\lambda=0}^{l_0} B_\lambda$ $(l_0<l)$ enthalten und $\widetilde{R}_1\cap C$ leer ist.

Wir betrachten nun denjenigen Teilraum T^* von T, der aus genau den Paaren $(\mathfrak{z}, t)$ besteht, wo $\mathfrak{z}\in E_t^1\cap(Z-\overline{Z}')$, $0\leq t\leq 1$. Die Überlagerung $(\Phi^{*-1}(T^*), \Phi^*, T^*)$ zerfällt in endlich viele Komponenten (B_μ^*, Φ^*, T^*), $\mu=0,\ldots,m$; es sei die Numerierung so gewählt, daß gilt: $B_0^*\subset R_0^*$. Der Raum B_0^* enthält sowohl Punkte $(r,0)$ als auch Punkte $(r,1)$. Jeder Punkt $(r,0)$ wird nun, da er wegen $B_0^*\subset R_0^*$ in $\widetilde{R}_0^*$ liegt, vermöge π in $\widetilde{R}_0\cap\Phi^{-1}(Z-\overline{Z}')$ und also in C abgebildet. Dann wird aber, da B_0^* zusammenhängend ist, sogar ganz B_0^* in C abgebildet. Andererseits wird aber jeder Punkt $(r,1)\in B_0^*$, da er in $\widetilde{R}_1^*$ liegt, durch π in $\widetilde{R}_1$ abgebildet. Da $\widetilde{R}_1\cap C$ leer ist, hat sich somit ein Widerspruch ergeben. Die Annahme, daß der Raum $\widetilde{R}$ nicht zusammenhängend ist, war also falsch: Hilfssatz 3 ist bewiesen.

Eine einfache Folgerung aus Hilfssatz 3 ist

Hilfssatz 4. Es sei $\hat{E}$ eine eindimensionale analytische Ebene des $(z_2,\ldots,z_n)$-Raumes durch $(z_2^{(0)},\ldots,z_n^{(0)})\in\hat{Z}$, die die Diskriminantenmenge $\hat{V}$ von V höchstens in $(z_2^{(0)},\ldots,z_n^{(0)})$ schneidet. Mit E^2 werde die zweidimensionale analytische Ebene $\{\mathfrak{z}|, (z_2,\ldots,z_n)\in\hat{E}\}$ bezeichnet. Dann ist $(\Phi^{-1}(E^2\cap Z-V), \Phi, E^2\cap Z-V)$ eine zusammenhängende analytisch-verzweigte Überlagerung von $E^2\cap Z-V$, die sich zu einer analytisch-verzweigten Überlagerung von $E^2\cap Z$ fortsetzen läßt.

Beweis. Wir haben nur zu zeigen, daß $(\Phi^{-1}(E^2\cap Z-V), \Phi, E^2\cap Z-V)$ zusammenhängend ist; die Fortsetzbarkeit ergibt sich direkt aus dem Fortsetzungssatz. Der Zusammenhang ist aber ebenfalls klar, denn nach Hilfssatz 3 ist stets $\Phi^{-1}(E_1^*\cap Z-V)$ zusammenhängend, wo E_1^* irgendeine eindimensionale analytische Ebene ist, die in der Form $\{\mathfrak{z}|, (z_2,\ldots,z_n)=(z_2^*,\ldots,z_n^*)\in\hat{E}-(z_2^{(0)},\ldots,z_n^{(0)})\}$ gegeben werden kann.

4. Unter Verwendung von Hilfssatz 2 und Hilfssatz 4 beweisen wir nun, indem wir die Terminologie des 3. Abschnittes beibehalten, die folgende Verallgemeinerung von Hilfssatz 2:

Hilfssatz 5. Es sei $p(r)$ eine plurisubharmonische Funktion in einer offenen Menge $R-D\supset\Phi^{-1}(Z-V)$, die in der Umgebung eines jeden Punktes von $D\subset\Phi^{-1}(V)$ nach oben beschränkt ist. Es sei $\widetilde{E}^1$ die eindimensionale analytische Ebene $\{z_2=z_2^{(0)},\ldots,z_n=z_n^{(0)}\}$. Bezeichnen $\widetilde{R}_1,\ldots,\widetilde{R}_k$ die endlich vielen

zusammenhängenden Komponenten von $\Phi^{-1}(\tilde{E}^1 - \mathfrak{z}_0)$, $\mathfrak{z}_0 = (z_1^{(0)}, \ldots, z_n^{(0)})$, *so gilt:*

$$\overline{\lim_{r \to r_0,\ r \in \tilde{R}_\varkappa}} p(r) = \overline{\lim_{r \to r_0,\ r \in R-D}} p(r) = m, \quad \varkappa = 1, \ldots, k.$$

Beweis. Wir zeigen zunächst, daß gilt:

$$\overline{\lim_{r \to r_0,\ r \in \tilde{R}_\varkappa}} p(r) = m_0 \quad \text{für jedes } \varkappa = 1, \ldots, k.$$

Zu dem Zwecke legen wir durch den Punkt $(z_2^{(0)}, \ldots, z_n^{(0)})$ eine eindimensionale analytische Ebene $\hat{E}^1$ in $\hat{Z}$, die $\hat{V}$ in der Nähe von $(z_2^{(0)}, \ldots, z_n^{(0)})$ nur in $(z_2^{(0)}, \ldots, z_n^{(0)})$ schneidet. Weiter denken wir uns zu $\hat{E}^1$ die zweidimensionale analytische Ebene E^2 wie in Hilfssatz 4 gewählt. Dann ist $(\Phi^{-1}(E^2 \cap Z - V), \Phi, E^2 \cap Z - V)$ nach Hilfssatz 4 zu einer *zusammenhängenden* Überlagerung von $E^2 \cap Z$ fortsetzbar. Aus Hilfssatz 2 ergibt sich alsdann unmittelbar die obige Behauptung.

Es bleibt zu zeigen, daß gilt: $m_0 = m$. Sicher gilt $m_0 \leq m$. Angenommen, es wäre $m_0 < m_1 < m$. Wir wählen dann einen Kreis

$$K: \{|z_1 - z_1^{(0)}| \leq \delta, \quad z_2 = z_2^{(0)}, \ldots, z_n = z_n^{(0)}\}, \quad \delta < d,$$

auf $\hat{E}^1$, der mit V nur den Punkt $\mathfrak{z}_0$ gemeinsam hat, derart, daß in $\Phi^{-1}(K - \mathfrak{z}_0)$ gilt: $p(r) < m_1$; letzteres ist offenbar möglich. Es läßt sich weiter eine Umgebung

$$W': \left\{\frac{\delta}{2} \leq |z_1 - z_1^{(0)}| \leq \delta, \quad |z_2 - z_2^{(0)}| < \varepsilon, \ldots, |z_n - z_n^{(0)}| < \varepsilon\right\}$$

finden, die keinen Punkt mit V gemeinsam hat, so daß für $r \in \Phi^{-1}(W')$ gilt: $p(r) < m_1$. Bezeichnen wir mit W die Umgebung

$$\{|z_1 - z_1^{(0)}| < \delta, \quad |z_\nu - z_\nu^{(0)}| < \varepsilon, \quad \nu = 2, \ldots, n\}$$

des Punktes $\mathfrak{z}_0$ und mit E irgendeine der Ebenen $\{z_2 = a_2, \ldots, z_n = a_n\}$, wo $|a_\nu - z_\nu^{(0)}| < \varepsilon$, $\nu = 2, \ldots, n$, so kann jede der Überlagerungen

$$(\Phi^{-1}(E \cap W - V), \Phi, E \cap W - V)$$

zu einer Überlagerung $(\tilde{R}, \Phi, E \cap W)$ von ganz $E \cap W$ fortgesetzt werden. Da $\tilde{R}$ stets eine RIEMANNsche Fläche ist, die durch eine holomorphe Abbildung τ spurpunkttreu in $\Phi^{-1}(E \cap W)$ abgebildet ist, kann man nach Satz 6 die Funktion p auf $\tilde{R} - \tau^{-1}(D)$ zu einer in ganz $\tilde{R}$ subharmonischen Funktion $\tilde{p}$ fortsetzen. Da wegen des Maximumprinzips die Funktion $\tilde{p}(r)$ im Innern von $\tilde{R}$ nicht größer als auf $\Phi^{-1}(\partial(E \cap W))$ sein kann, so gilt für jeden Punkt $r \in \tilde{R}$: $\tilde{p}(r) < m_1$. Es ist deshalb $p(r) \leq m_1$ in jedem Punkt von $\Phi^{-1}(W) - D$. Dann gilt aber auch $\overline{\lim_{r \to r_0,\ r \in R-D}} p(r) \leq m_1 < m$ im Widerspruch zur Voraussetzung. Hilfssatz 5 ist bewiesen.

5. Ist $a(r)$ irgendeine nach oben halbstetige Funktion auf einer s-blättrigen, analytisch-verzweigten, zusammenhängenden Überlagerung (R, Φ, G), so ordnen wir derselben eine in G nach oben halbstetige Funktion $b(\mathfrak{z})$ zu,

indem wir setzen: $b(\mathfrak{z}) = \sum_{\sigma=1}^{s} a(r_\sigma)$, wo r_σ die Punkte aus $\Phi^{-1}(\mathfrak{z})$ durchläuft. Dabei komme der Punkt r_σ genau t-mal vor, wenn die Überlagerung R in hinreichend kleiner Umgebung von r_σ stets t Blätter hat. *Wir nennen* $b(\mathfrak{z})$ *die zu* $a(r)$ *gehörende elementarsymmetrische Funktion* und beweisen:

Hilfssatz 6. Es sei $\breve{p}(r)$ *eine in* R *nach oben halbstetige und in* $R-D$, $D \subset \Phi^{-1}(V)$, *plurisubharmonische Funktion; für jeden Punkt* $r_0 \in D$ *gelte:* $\overline{\lim}_{r \to r_0,\, r \in R-D} \breve{p}(r) = \breve{p}(r_0)$. *Dann ist die zu* $\breve{p}(r)$ *gehörende elementarsymmetrische Funktion* $\breve{q}(\mathfrak{z})$ *plurisubharmonisch in* G.

Beweis. Sei zunächst R eine Überlagerung eines Gebietes G der eindimensionalen Zahlenebene, also eine RIEMANNsche Fläche. Dann besteht D aus isolierten Punkten; da die Aussage des Hilfssatzes lokaler Natur ist, dürfen wir annehmen, daß G ein Kreis und D über dem Mittelpunkt $z^{(0)}$ dieses Kreises liegt und daß R zusammenhängend und nur über $z^{(0)}$ verzweigt ist. Da R im vorliegenden Falle eine RIEMANNsche Fläche ist, folgt aus Satz 5, daß $\breve{p}$ in ganz R subharmonisch ist. Es seien nun $\lambda_1, \ldots, \lambda_s$ die Decktransformationen von R. Da dieselben holomorphe Abbildungen sind, folgt unter Verwendung der Eigenschaften 1) und g) aus § 1, daß auch die Funktion $\breve{q}'(r) = \sum_{\sigma=1}^{s} \breve{p} \circ \lambda_\sigma(r)$ subharmonisch in R ist. Offenbar hat $\breve{q}'(r)$ in allen Punkten von R, die über demselben Grundpunkt $z \in G$ liegen, den gleichen Wert. Da überdies $\breve{q}(z) = \breve{q}'\big(\Phi^{-1}(z)\big)$, so ist also $\breve{q}(z)$ subharmonisch in G.

Sei nun (R, Φ, G) irgendeine Überlagerung. Da die Summe plurisubharmonischer Funktionen plurisubharmonisch ist, ist zunächst $q(\mathfrak{z}) = \sum_{1}^{s} \breve{p}(r_\sigma)$, $\mathfrak{z} \in G - V$, eine in $G - V$ plurisubharmonische Funktion. Da $q(\mathfrak{z})$ in der Umgebung eines jeden Punktes von V nach oben beschränkt ist, läßt sich $q(\mathfrak{z})$ nach Satz 6 zu einer in ganz G plurisubharmonischen Funktion $\tilde{q}(\mathfrak{z})$ fortsetzen. Wir behaupten: $\tilde{q}(\mathfrak{z}) = \breve{q}(\mathfrak{z})$. Zu dem Zwecke genügt es zu zeigen:

$$(*) \qquad \overline{\lim_{\mathfrak{z} \to \mathfrak{z}_0,\, \mathfrak{z} \in G-V}} q(\mathfrak{z}) \geq \breve{q}(\mathfrak{z}_0) \quad \text{für jedes } \mathfrak{z}_0 \in V.$$

Sei daher $\mathfrak{z}_0 \in V$ irgendein Punkt. Man kann um $\mathfrak{z}_0$ einen in G enthaltenen Polyzylinder Z legen, der in bezug auf V die im Hilfssatz 3 geforderten Eigenschaften hat (eventuell ist eine lineare Koordinatentransformation auszuführen). Ohne Einschränkung der Allgemeinheit werde angenommen, daß über $\mathfrak{z}_0$ nur ein Punkt r_0 von R liegt. Es sei wieder die Ebene $\tilde{E}^1$ wie im Hilfssatz 5 gewählt; es bezeichne $(\tilde{R}, \tilde{\Phi}, \tilde{E}^1 \cap Z)$ die Fortsetzung der Überlagerung $\big(\Phi^{-1}(\tilde{E}^1 \cap Z - V), \Phi, \tilde{E}^1 \cap Z - V\big)$. Zerfällt nun der Raum $\tilde{R}$ etwa in k zusammenhängende Komponenten $\tilde{R}_\varkappa$, $\varkappa = 1, \ldots, k$, so setzen wir zunächst die Funktion $\breve{p}$ aus $\Phi^{-1}(\tilde{E}^1 \cap Z - V)$ in jede Komponente $\tilde{R}_\varkappa$ zu einer dort subharmonischen Funktion $\breve{p}_\varkappa$ fort. Unter Verwendung von Hilfssatz 5 folgt: $\breve{p}_\varkappa(r) = \breve{p}(r_0)$ für $r \in \tilde{\Phi}^{-1}(\mathfrak{z}_0)$. Daher gilt auf $\tilde{E}^1 \cap Z$, wenn $\breve{q}_\varkappa$ die zu $\breve{p}_\varkappa$

gehörende elementarsymmetrische Funktion bezeichnet:

$$\breve{q}(\mathfrak{z}) = \sum_{\varkappa=1}^{k} \breve{q}_\varkappa(\mathfrak{z}).$$

Da nach dem bereits Bewiesenen die Funktionen $\breve{q}_\varkappa(\mathfrak{z})$ sämtlich subharmonisch sind, ist also auch $\breve{q}(\mathfrak{z})$ subharmonisch auf $\tilde{E}^1 \cap Z$. Es gilt daher:

$$\overline{\lim_{\mathfrak{z}\to\mathfrak{z}_0,\ \mathfrak{z}\in\tilde{E}^1\cap Z-\mathfrak{z}_0}} q(\mathfrak{z}) = \breve{q}(\mathfrak{z}_0).$$

Dann gilt aber erst recht die Ungleichung (*), w. z. b. w.

6. Wir sind nun in der Lage, Satz 3 für beliebige komplexe Räume X zu beweisen. Da die Aussage des Satzes lokaler Natur ist, dürfen wir ohne Einschränkung der Allgemeinheit voraussetzen, daß der komplexe Raum X eine endlich-blättrige, analytisch-verzweigte Überlagerung (R, Φ, G) eines Gebietes G des C^n ist. Sei also D eine dünne Menge der Ordnung 1 in R und p eine in $R-D$ plurisubharmonische Funktion, die in der Umgebung eines jeden Punktes $r \in D$ nach oben beschränkt ist. Da wir p nach Satz 6 in jeden uniformisierbaren Punkt von R eindeutig plurisubharmonisch fortsetzen können, dürfen wir annehmen, daß D in der Menge $\Phi^{-1}(V)$ enthalten ist.

Wir zeigen zunächst, daß höchstens eine plurisubharmonische Fortsetzung von p aus $R-\Phi^{-1}(V)$ in R möglich ist. Angenommen, es gäbe zwei solche Fortsetzungen $\breve{p}$ und $\breve{p}'$. Dann gilt also $\breve{p}'(r) = \breve{p}(r)$ für jeden Punkt $r \in R - \Phi^{-1}(V)$. Ist nun $r_0 \in \Phi^{-1}(V)$ irgendein Punkt, so legen wir eine eindimensionale analytische Ebene E durch $\mathfrak{z}_0 = \Phi(r_0)$, derart, daß E in der Nähe von $\mathfrak{z}_0$ die Menge V nur in $\mathfrak{z}_0$ schneidet. Es zerfällt $\Phi^{-1}(E-\mathfrak{z}_0)$ in der Umgebung von $\mathfrak{z}_0$ in endlich viele zusammenhängende RIEMANNsche Flächen; der Punkt r_0 ist Häufungspunkt wenigstens einer derselben, etwa $\tilde{F}$. Dann folgt, da $\breve{p}$ und $\breve{p}'$ nach Voraussetzung beide plurisubharmonisch in R sind und in $R-\Phi^{-1}(V)$ mit p übereinstimmen, unter Verwendung von Hilfssatz 5:

$$\overline{\lim_{r\to r_0,\ r\in\tilde{F}}} p(r) = \begin{cases} \breve{p}(r_0), \\ \breve{p}'(r_0). \end{cases}$$

Daher gilt auch $\breve{p}(r_0) = \breve{p}'(r_0)$; mithin gibt es höchstens eine plurisubharmonische Fortsetzung von p in R.

Wir werden nun zeigen, daß die Funktion

$$\breve{p}(r) = \overline{\lim_{r'\to r,\ r'\in R-D}} p(r'), \qquad D < \Phi^{-1}(V),$$

tatsächlich eine plurisubharmonische Fortsetzung von p in R ist.

Zu dem Zwecke sei τ irgendeine holomorphe Abbildung eines Gebietes W der z-Ebene in R; wir dürfen annehmen, daß τ nicht konstant ist. Dann ist die zur Abbildung

$$\Phi \circ \tau: \quad z_1 = f_1(z), \ldots, z_n = f_n(z)$$

gehörende Funktionalmatrix $(f_1'(z), \ldots, f_n'(z))$ bis auf eine in W isoliert liegende Punktmenge S' vom Maximalrang. Es gibt deshalb zu jedem Punkt $z^{(0)} \in W - S'$ eine Umgebung $U'(z^{(0)}) < W - S'$, derart, daß $\Phi \circ \tau(U')$ eine in einer Umgebung U'' von $\zeta_0 = \Phi \circ \tau(z^{(0)})$ eindimensionale analytische Menge A ist. Wir ziehen nun folgenden hier nicht zu beweisenden Hilfssatz heran (zum Beweise desselben vgl. [6]).

Hilfssatz 7. Es sei (R'', Φ, U'') eine analytisch-verzweigte Überlagerung; es sei A eine eindimensionale analytische Menge in U''. Dann gibt es eine isolierte Punktmenge S^* in A, so daß $(\Phi^{-1}(A - S^*), \Phi, A - S^*)$ eine unverzweigte Überlagerung ist.

Auf Grund dieses Hilfssatzes gibt es zu jedem Punkt $\zeta_1 \in A - S^*$ eine n-dimensionale Umgebung $U^*(\zeta_1) < U''$, derart, daß jede zusammenhängende Komponente $R_\varkappa$ von $\Phi^{-1}(U^*)$ über $(A - S^*) \cap U^*$ einblättrig ist. Ist nun $\breve{p}_\varkappa$ die Beschränkung von $\breve{p}$ auf $R_\varkappa$ und bezeichnet $\breve{q}_\varkappa$ die in U^* definierte, zu $\breve{p}_\varkappa$ gehörende elementarsymmetrische Funktion, so ist nach Hilfssatz 6 jede dieser Funktionen $\breve{q}_\varkappa$ plurisubharmonisch in U^*. Da über jedem Punkt von $(A - S^*) \cap U^*$ jeweils nur ein Punkt von $R_\varkappa$ liegt, stimmt $\breve{p}_\varkappa$ auf

$$\Phi^{-1}\big((A - S^*) \cap U^*\big) \cap R_\varkappa$$

mit $\breve{q}_\varkappa$ auf $(A - S^*) \cap U^*$ überein. Daraus folgt aber, daß die Beschränkung von $\breve{p}$ auf $\Phi^{-1}(A - S^*)$ eine subharmonische Funktion ist. Setzt man nun $S = (\Phi \circ \tau)^{-1}(S^*) \cap U'(z^{(0)})$, so ist S eine isolierte Punktmenge in $U'(z^{(0)})$; da τ die Menge $U' - S$ in $\Phi^{-1}(A - S^*)$ abbildet, ist $\breve{p} \circ \tau$ subharmonisch in $U' - S$. Es bleibt zu zeigen, daß $\breve{p} \circ \tau$ auch in allen Punkten von S selbst subharmonisch ist; sei also $z^{(1)} \in S$ irgendein Punkt. Wir wählen eine Umgebung $\hat{U}\big(\Phi \circ \tau(z^{(1)})\big)$ so, daß $\big(\Phi^{-1}(\hat{U}), \Phi, \hat{U}\big)$ eine zusammenhängende, s-blättrige Komponente $\hat{R}$ enthält, die über $\zeta_1 = \Phi \circ \tau(z^{(1)})$ nur den einen Punkt $r_1 = \tau(z^{(1)})$ besitzt.

Es sei $\breve{p}(r_1) = m$. Wir bilden in $\hat{U}$ zu der Beschränkung von $\breve{p}$ auf $\hat{R}$ die elementarsymmetrische Funktion $\breve{q}(\zeta)$; dieselbe ist nach Hilfssatz 6 in $\hat{U}$ plurisubharmonisch. Deshalb ist $\breve{q} \circ \Phi \circ \tau$ in der Umgebung von $z^{(1)}$ subharmonisch. Es gilt: $\overline{\lim\limits_{z \to z^{(1)}}} \frac{1}{s} \breve{q} \circ \Phi \circ \tau(z) = m$.

Andererseits gilt, falls man $\overline{\lim\limits_{z \to z^{(1)}}} \breve{p} \circ \tau(z) \leq m_1 < m$ annimmt:

$$\frac{1}{s} \overline{\lim_{z \to z^{(1)}}} \breve{q} \circ \Phi \circ \tau(z) = \overline{\lim_{z \to z^{(1)}}} \frac{1}{s} \Big(\breve{p} \circ \tau(z) + \sum_{\substack{r_\sigma \in \Phi^{-1}\Phi\tau(z) \cap \hat{R} \\ r_\sigma \neq \tau(z)}} \breve{p}(r_\sigma) \Big)$$

$$\leq \frac{m_1}{s} + \frac{s-1}{s} m < m.$$

Also muß gelten $\overline{\lim\limits_{z \to z^{(1)}}} \breve{p} \circ \tau(z) = m$; d.h. aber, daß $\breve{p} \circ \tau(z^{(1)})$ mit der subharmonischen Fortsetzung von $\breve{p} \circ \tau$ in $z^{(1)}$ übereinstimmt. Die Funktion $\breve{p} \circ \tau$ ist somit in jedem Punkt von U' subharmonisch. Daraus folgt, daß $\breve{p} \circ \tau$ in $W - S'$ subharmonisch ist. Wiederholt man den soeben durchgeführten Schluß für die ebenfalls isoliert liegenden Punkte von S', so ergibt sich, daß $\breve{p} \circ \tau$ überall in W subharmonisch ist. Damit ist Satz 3 vollständig bewiesen.

7. Wir beweisen nun abschließend Satz 4, indem wir zeigen, daß unter seinen Voraussetzungen auch die Voraussetzungen von Satz 3 erfüllt sind. Wir können wieder annehmen, daß X eine analytisch-verzweigte Überlagerung (R, Φ, G) ist und dürfen weiter, falls G hinreichend klein gewählt ist, laut Definition der dünnen Menge sogar voraussetzen, daß es eine in G analytische, höchstens $(n-2)$-dimensionale Menge M gibt, so daß gilt: $D \subset \Phi^{-1}(M)$. Wir zeigen, daß $p(r)$ in einer Umgebung eines jeden Punktes aus D nach oben beschränkt ist. Damit ist dann der Satz bewiesen, da eine dünne Menge der Ordnung 2 erst recht dünn von der Ordnung 1 ist. Angenommen, $r' \in D$ sei ein Punkt, in dessen Umgebung p nicht nach oben beschränkt ist. Dann gibt es eine gegen r' konvergierende Punktfolge $r_\nu \in R - \Phi^{-1}(M \cup V)$[7]), so daß gilt: $\lim_{\nu\to\infty} p(r_\nu) = +\infty$. Wir setzen $z_\nu = \Phi(r_\nu)$, $z' = \Phi(r')$; dann gilt $\lim_{\nu\to\infty} z_\nu = z'$. Es sei nun E' eine eindimensionale analytische Ebene durch z', die $M \cup V$ in einer Hyperkugel $H(z') \subset\subset G$ nur in z' schneidet. Es gibt dann eine gegen E' konvergierende Folge E_ν von eindimensionalen analytischen Ebenen durch z_ν, die $M \cap H$ nirgends und $V \cap H$ nur punkthaft schneiden. Es gibt weiter eine reelle Zahl k, so daß für jeden Punkt $r \in \Phi^{-1}(E' \cap \partial H)$ gilt: $p(r) < k$. Aus der Halbstetigkeit von p nach oben folgt, daß diese Ungleichung auch noch in einer ganzen Umgebung U von $\Phi^{-1}(E' \cap \partial H)$ gilt. Wählt man ν hinreichend groß, so ist $\Phi^{-1}(E_\nu \cap \partial H)$ in U enthalten. Da $\left(\Phi^{-1}(E_\nu \cap H - V), \Phi, E_\nu \cap H - V\right)$ zu einer Überlagerung von $E_\nu \cap H$ fortsetzbar ist, ergibt das Maximumprinzip, daß $p(r)$ auf $\Phi^{-1}(E_\nu \cap H)$ kleiner als k ist. Deshalb gilt auch $p(r_\nu) \leq k$, da $r_\nu \in \Phi^{-1}(E_\nu \cap H)$. Daraus folgt aber $\lim_{\nu\to\infty} p(r_\nu) \leq k$ im Widerspruch zur Annahme. Satz 4 ist bewiesen.

Literatur

[1] BREIOT, M.: Étude des fonctions sousharmoniques au voisinage d'un point. Act. scient. et ind. **139**, II, 5—55 (1934). — [2] BREMERMANN, H.: On the conjecture of the equivalence of the plurisubharmonic functions and the Hartogs functions. Math. Ann. **131**, 76—86 (1956). — [3] CARTAN, H.: Séminaire E.N.S. Paris 1951/52 u. 1953/54 (hektographiert). — [4] GRAUERT, H., u. R. REMMERT: Zur Theorie der Modifikationen I. Math. Ann. **129**, 274—296 (1955). — [5] GRAUERT, H., et R. REMMERT: Fonctions plurisousharmoniques dans des espaces analytiques. Généralisation d'un théorème d'Oka. C. R. Acad. Sci. Paris **241** 1371—1373 (1955). — [6] GRAUERT, H., u. R. REMMERT: Analytisch-verzweigte Überlagerungen und komplexe Räume. Erscheint demnächst in den Math. Ann. — [7] HIRZEBRUCH, F.: Über vierdimensionale Riemannsche Flächen mehrdeutiger analytischer Funktionen von zwei komplexen Veränderlichen. Math. Ann. **126**, 1—22 (1953). — [8] HITOTUMATU, S.: On some conjectures concerning pseudo-convex domains. J. Math. Soc. Jap. **6**, 173—195 (1954). — [9] LELONG, P.: Fonctions plurisousharmoniques; mesures de Radon associées. Applications aux fonctions analytiques. Colloque sur les Fonctions de plusieurs variables, Bruxelles, 21—40, 1953. — [10] OKA, K.: Sur les fonctions analytiques de plusieurs variables. IX: Domaines finis sans point critique intérieur. Jap. J. Math. **23**, 97—155 (1953). — [11] REMMERT, R., u. K. STEIN: Über die wesentlichen Singularitäten analytischer Mengen. Math. Ann. **126**, 263—306 (1953).

[7]) Das folgt unter Verwendung von Satz 3!

Münster i. W., Mathematisches Institut der Universität

(Eingegangen am 1. Januar 1956)

Bojanić, R., W. Jurkat u. A. Peyerimhoff
Math. Zeitschr. Bd. 65, S. 195—199 (1956)

Über einen Taubersatz für Faltungen

Von
R. Bojanić, W. Jurkat und A. Peyerimhoff

Von Herrn Erdös wurde kürzlich [3] die Frage aufgeworfen nach Aussagen über die Größenordnung von Folgen $\{s_n\}$ mit den Eigenschaften

$$\sum_{\nu=0}^{n} s_{n-\nu} a_\nu = \frac{n^2}{2} + O(n) \qquad \left(s_n = \sum_{\nu=0}^{n} a_\nu\right) \tag{1}$$

und

$$a_n \geq 0. \tag{2}$$

Herr Erdös zeigt, daß stets $s_n = n + o(n)$ ist, während $s_n = n + o(\sqrt{n})$ nicht gefolgert werden kann.

Herr Avakumović hat bewiesen [1], daß aus (1) und (2) die Beziehung $s_n = n + O\left(n^{\sqrt{3}-1+\varepsilon}\right)$ folgt für jedes $\varepsilon > 0$.

Wir werden im folgenden zeigen, daß sogar $s_n = n + O\left(n^{\frac{1}{3}} (\log n)^{\frac{1}{3}}\right)$ aus (1) und (2) gefolgert werden kann. Es verlangt keine Mehrarbeit, dieses Ergebnis gleich in etwas allgemeinerer Form für Stieltjes-Integrale auszusprechen und zu beweisen.

§ 1. Es sei $A(t)$ eine Funktion mit den Eigenschaften

$$A(0) = 0, \quad A(t_2) \geq A(t_1) \quad \text{für } t_2 \geq t_1 \geq 0 \tag{3}$$

und

$$A(t) = \tfrac{1}{2}\big(A(t+0) + A(t-0)\big) \quad \text{für alle } t > 0.$$

Das Integral $h(x) = \int_0^x A(x-t)\, dA(t)$ existiert für alle $x \geq 0$ mit eventueller Ausnahme einer abzählbaren Punktmenge und kann überall so festgesetzt werden, daß $h(x)$ eine monoton wachsende Funktion ist mit $h(0) = 0$, $h(x) = \tfrac{1}{2}\big(h(x+0) + h(x-0)\big)$ für alle $x > 0$ (vgl. [4], theorem 11.2a, S. 84 und den Beweis auf S. 85). Für jede Funktion $A(t)$ mit der Eigenschaft (3) ist damit das Integral $\int_0^x A(x-t)\, dA(t)$ eindeutig erklärt. Wir verwenden bei unseren Überlegungen den folgenden einseitigen Konvexitätssatz.

Hilfssatz. Es sei $k > 0$ und $W(x)$ sei eine Funktion mit den folgenden Eigenschaften:

$$\begin{cases} W(x) \text{ ist erklärt (und endlich) für } x \geq 0, \\ 0 < W(x_1) \leq W(x_2) \text{ für } 0 \leq x_1 \leq x_2, \\ \text{mit positiven Konstanten } C_1 \text{ und } C_2 \text{ ist } W(x') \leq C_1 W(x) \\ \text{für } 0 \leq x < x' \leq x + C_2 \big(W(x)\big)^{1/(k+1)}. \end{cases} \tag{4}$$

Ist ferner $C(t)$ für $t \geq 0$ eine meßbare Funktion, die in jedem Intervall $0 \leq t \leq T$ beschränkt ist, so folgt aus

$$\int_0^x (x-t)^{k-1} C(t)\, dt = O\big(W(x)\big) \quad (x \to \infty) \tag{5}$$

und

$$\begin{cases} C(x+t) - C(x) \geq C_3 (W(x))^{\frac{1}{k+1}}, \quad C(x) - C(x-t) \geq C_3 (W(x))^{\frac{1}{k+1}} \\ (C_3 \text{ eine beliebige Konstante}) \text{ für alle } x \geq 0 \text{ und} \\ 0 < t \leq C_4 (W(x))^{1/(k+1)} \; (C(y) = 0 \text{ für } y < 0) \end{cases} \tag{6}$$

die Beziehung

$$C(x) = O(1)\, (W(x))^{\frac{1}{k+1}} \quad (x \to \infty). \tag{7}$$

Für einen Beweis vgl. [2]; § 1.8, S. 23 (die Beweismethode bleibt bei den hier vorgenommenen Änderungen anwendbar). Im folgenden werden wir nur den Spezialfall $k=2$, $W(x) = x^2 \log x$ $(x \geq 2)$ verwenden.

Es gilt nun der

Satz 1. *Ist $A(t)$ eine Funktion mit den Eigenschaften (3), so folgt aus*

$$h(x) = \int_0^x A(x-t)\, dA(t) = \frac{x^2}{2} + O(x) \quad (x \to \infty) \tag{8}$$

die Beziehung

$$A(x) = x + O(1)\, x^{\frac{1}{2}} (\log x)^{\frac{1}{2}} \quad (x \to \infty). \tag{9}$$

Wir zerlegen den Beweis in einzelne Teile.

1. Wegen (3) und (8) ist

$$A^2\Big(\frac{x}{2}\Big) \leq \int_0^{x/2} A(x-t)\, dA(t) \leq \int_0^x A(x-t)\, dA(t) = O(x^2) \quad (x \to \infty)$$

für alle x mit eventueller Ausnahme einer abzählbaren Punktmenge. Wegen (3) ist somit für alle x

$$A(x) = O(x) \quad (x \to \infty). \tag{10}$$

2. Wegen (3) und (10) ist $f(s) = \int_0^\infty e^{-st}\, dA(t)$ konvergent für $\Re s > 0$. Für $\Re s > 0$ ist also $f^2(s) = \int_0^\infty e^{-st}\, dh(t) = s \int_0^\infty e^{-st} h(t)\, dt$ (vgl. [4], theorem 11.7b, S. 91 und beachte, daß $h(0) = 0$ ist). Wegen (8) ist nun

$$f^2(s) = \frac{1}{s^2}\Big(1 + g(s)\, \frac{s^3}{\sigma^2}\Big) \quad \text{mit} \quad |g(s)| \leq K \text{ für } 0 < \sigma = \Re s \leq 1. \tag{11}$$

3. Im folgenden sei stets $s = \sigma + it$. Für $\frac{|s|^3}{\sigma^2} < \frac{1}{2K}$ und $0 < \sigma \leq 1$ folgt aus (11) die Relation

$$f(s) = \frac{1}{s}\Big(1 + g_1(s)\, \frac{s^3}{\sigma^2}\Big) \quad \text{mit} \quad |g_1(s)| \leq K_1. \tag{12}$$

[Das Vorzeichen dieser Wurzel ergibt sich aus der Tatsache, daß $f(s) > 0$ ist für $s > 0$.]

Für $\frac{|s|^3}{\sigma^2} \geq \frac{1}{2K}$ und $0<\sigma\leq 1$ folgt aus (11) die Abschätzung

$$|f(s)| \leq \frac{1}{|s|}\left(1+|g(s)|\frac{|s|^3}{\sigma^2}\right)^{\frac{1}{3}} \leq K_2 \frac{|s|^{\frac{1}{3}}}{\sigma}. \tag{13}$$

4. Wegen (3) $(A(0)=0)$ ist $f(s)-\frac{1}{s} = s\int_0^\infty (A(t)-t)e^{-st}dt$ für $\sigma>0$. Setzt man zur Abkürzung $B(t)=\int_0^t (t-\tau)(A(\tau)-\tau)\,d\tau$ und $\frac{1}{s^3}\left(f(s)-\frac{1}{s}\right)=\eta(s)$, so ist (vgl. [4], theorem 12.1a, S. 91)

$$\eta(s) = \int_0^\infty e^{-st} B(t)\,dt, \tag{14}$$

wobei dieses Integral absolut konvergent ist für $\sigma>0$. Die Funktion $\eta(s)$ erfüllt für $0<\sigma\leq 1$ die Beziehung

$$|\eta(s)| \leq \begin{cases} \frac{K_1}{|s|\,\sigma^2} & \text{für} \quad |s|^3 < \frac{1}{2K}\sigma^2 \\ \frac{K_3}{|s|^{\frac{1}{3}}\sigma} & \text{für} \quad |s|^3 \geq \frac{1}{2K}\sigma^2. \end{cases} \tag{15}$$

Die erste dieser Beziehungen folgt unmittelbar aus (12), die zweite aus (13) zusammen mit der aus $|s|^3 \geq \frac{1}{2K}\sigma^2$ folgenden Ungleichung $\frac{1}{|s|^4} \leq (2K)^{\frac{1}{3}} \frac{1}{\sigma |s|^{\frac{1}{3}}}$.

5. Aus (14) folgt die Beziehung

$$B(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \eta(s)\, e^{xs}\,ds \quad (x>0), \tag{16}$$

wobei das Integral wegen (15) absolut konvergiert für $c>0$ (vgl. [4], theorem 7.3, S. 66). Mit $0<\sigma=c\leq 1$ folgt aus (16) wegen (15) die Abschätzung

$$e^{-xc}|B(x)| \leq \frac{K_4}{c^2}\int_{|s|^3\leq \frac{1}{2K}\sigma^3} \frac{dt}{|c+it|} + \frac{K_5}{c}\int_{|s|^3\geq \frac{1}{2K}\sigma^3} \frac{dt}{|c+it|^{\frac{1}{3}}} = \mathrm{I}+\mathrm{II}.$$

Setzt man $c=\frac{1}{x}$, so ist für $x\to\infty$ und passende Konstanten $\varrho>0$, $\delta>0$,

$$\mathrm{I} = O(x^2)\int_0^{x^{-1}} \frac{dt}{x^{-1}} + O(x^2)\int_{x^{-1}}^{\varrho x^{-\frac{1}{3}}} \frac{dt}{t} = O(x^2\log x)$$

und

$$\mathrm{II} = O(x)\int_{\delta x^{-\frac{1}{3}}}^\infty \frac{dt}{t^{\frac{1}{3}}} = O(x^2).$$

Zusammengefaßt ergibt sich also

$$B(x) = O(x^2\log x) \quad (x\to\infty). \tag{17}$$

6. Setzen wir im Hilfssatz $k=2$, $W(x)=x^2\log x$ $(x\geq 2)$, $W(x)=4\log 2$ $(0\leq x\leq 2)$ [diese Funktion erfüllt die Bedingung (4)] und $C(t)=A(t)-t$, so ist (5) erfüllt wegen (17). Aus (3) folgen die Beziehungen $C(x+t)-C(x)\geq -t$, $C(x)-C(x-t)\geq -t$, so daß auch (6) erfüllt ist. Die Behauptung (7) des

Hilfssatzes ist nun gleichbedeutend mit der zu beweisenden Beziehung (9). Damit ist Satz 1 bewiesen.

Ist $\{a_\nu\}$ eine Folge nicht negativer Zahlen und setzen wir $A(0)=0$, $A(x)=\sum_{\nu<x} a_\nu$ für $x\neq 0, 1, 2, \ldots$, $A(n)=\frac{1}{2}\big(A(n+0)+A(n-0)\big)$, so erfüllt diese Funktion die Bedingungen (3). Erfüllt $\{a_\nu\}$ noch die Bedingung (1), so genügt $A(x)$ den Voraussetzungen von Satz 1, denn es ist für $n<x<n+1$

$$h(x)=\int_0^x A(x-t)\,dA(t)=\sum_{\nu=0}^{n} s_{n-\nu}a_\nu=\frac{x^2}{2}+O(x)\quad(x\to\infty).$$

Als Anwendung von Satz 1 erhalten wir so den zu Anfang dieser Arbeit ausgesprochenen Sachverhalt.

Satz 2. *Erfüllt eine Folge $\{a_n\}$ die Bedingungen (1) und (2), so ist $s_n=n+O(1)\,n^{\frac{3}{4}}(\log n)^{\frac{1}{4}}$ $(n\to\infty)$.*

§ 2. Man wird bei einer genauen Durchsicht des Beweises von Satz 1 feststellen, daß die Voraussetzung $A(t_1)\leq A(t_2)$ $(0\leq t_1\leq t_2)$ im wesentlichen nur im Punkt 1 verwendet wird (d.h. bei der Feststellung der Konvergenz von $\int_0^\infty e^{-st}\,dA(t)$ für $\Re s>0$), während etwa bei der Anwendung des Konvexitätssatzes die allgemeinere Voraussetzung $A(t_2)-A(t_1)\geq -K(t_2-t_1+1)$ $(t_2\geq t_1)$ ausreichend gewesen wäre. So erhebt sich die Frage, ob Satz 1 unter dieser allgemeineren Voraussetzung noch richtig bleibt, wobei die Behauptung natürlich dahingehend zu erweitern ist, daß eine der beiden Möglichkeiten $A(x)=\pm x+O(1)\,x^{\frac{3}{4}}(\log x)^{\frac{1}{4}}$ eintritt. Das folgende Beispiel lehrt, daß dies nicht zutrifft. Wir zeigen, daß Satz 2 mit $a_n\geq -K$ statt (2) falsch wird.

Es sei $\{a_n\}$ erklärt durch $\sum a_n z^n=-\frac{1}{2}\cdot\frac{\sqrt{1-2z}\sqrt{1-5z}}{1-z}$. Setzt man

$$b_n=-\sum_{\nu=0}^{n} A_{n-\nu}^{-\frac{3}{2}}A_\nu^{-\frac{3}{2}}\left(\frac{2}{5}\right)^\nu\qquad\left(A_n^k=\binom{n+k}{n}\right),$$

so gilt

$$\frac{b_n}{A_n^{-\frac{3}{2}}}=\frac{-1}{A_n^{-\frac{3}{2}}}\sum_{\nu<\frac{1}{2}n}A_{n-\nu}^{-\frac{3}{2}}A_\nu^{-\frac{3}{2}}\left(\frac{2}{5}\right)^\nu+O\left(\left(\frac{2}{5}\right)^{\frac{1}{2}n}n^{\frac{3}{2}}\right)$$

$$\to-\sum_{\nu=0}^{\infty}A_\nu^{-\frac{3}{2}}\left(\frac{2}{5}\right)^\nu=-\sqrt{\frac{3}{5}}\quad(n\to\infty).$$

Wegen $a_n\asymp\frac{1}{2}\sum_{\nu=0}^{n}5^\nu b_\nu$ ist somit $a_n\geq C\frac{5^n}{n^{\frac{3}{2}}}$ für ein $C>0$ und alle großen n, so daß $a_n\geq -K$ ist, während $s_n\asymp\pm n$ nicht richtig ist. Mit der Abkürzung $t_n=\sum_{\nu=0}^{n}s_{n-\nu}a_\nu$ ist schließlich $\sum t_n z^n=\frac{1}{4}\cdot\frac{(1-2z)(1-5z)}{(1-z)^3}$, woraus sich sofort die Beziehung $t_n=\frac{n^2}{2}+O(n)$ $(n\to\infty)$ ergibt.

Der Satz 1 bleibt aber richtig, wenn wir in (3) die Bedingung $A(t_1)\leq A(t_2)$ $(0\leq t_1\leq t_2)$ ersetzen durch $A(t_2)-A(t_1)\geq -K(t_2-t_1+1)$ $(0\leq t_1\leq t_2)$[1] und

[1] Außerdem ist stillschweigend vorausgesetzt, daß $A(t)$ von beschränkter Schwankung ist in jedem Intervall $0\leq t\leq T$.

zusätzlich noch fordern, daß das Integral $\int_0^\infty e^{-st}\,dA(t)$ für $\Re s > 0$ konvergiert. Durch die Forderung $h(0)=0$, $h(x)=\frac{1}{2}\big(h(x+0)+h(x-0)\big)$ $(x>0)$ ist das Integral $h(x)=\int_0^x A(x-t)\,dA(t)$ für alle $x \geqq 0$ in eindeutiger Weise erklärt (vgl. [4], S. 84), und es gilt die folgende Ergänzung zu Satz 1.

Satz 3. *Es sei $A(t)$ eine Funktion mit den folgenden Eigenschaften*[1])

$$(18)\qquad \begin{cases} A(0)=0, \quad A(t_2)-A(t_1) \geqq -K(t_2-t_1+1) & \textit{bzw.} \\ A(t_2)-A(t_1) \leqq K(t_2-t_1+1) & \textit{für} \quad 0 \leqq t_1 \leqq t_2, \\ A(t)=\frac{1}{2}\big(A(t+0)+A(t-0)\big) \quad (t>0). \end{cases}$$

Erfüllt $A(t)$ die Bedingungen (8) *und*

$$(19)\qquad \overline{\lim_{t\to\infty}} \frac{\log|A(t)|}{t} \leqq 0,$$

so gilt eine der Beziehungen

$$(20)\qquad A(x)=\pm x + O(1)\,x^{\frac{1}{3}}(\log x)^{\frac{1}{3}} \quad (x\to\infty).$$

Der Beweis verläuft fast ebenso wie der Beweis von Satz 1. Der Schritt 1. fällt weg, die Existenz von $f(s)=\int_0^\infty e^{-st}\,dA(t)$ für $\Re s>0$ ergibt sich aus (19) und die Überlegungen von Schritt 2. bleiben ungeändert, insbesondere ist (11) richtig. Für $\frac{|s|^3}{\sigma^2} < \frac{1}{2K}$, $0<\sigma\leqq 1$ besitzt $f(s)$ wegen (11) keine Nullstellen und wir dürfen ohne Beschränkung der Allgemeinheit annehmen, daß $f(s)>0$ ist für kleine $s>0$ [beachte (18)]. Die folgenden Schritte des Beweises von Satz 1 bleiben ungeändert bis auf die Anwendung des Hilfssatzes in 6., wo aber unsere allgemeinere Voraussetzung (18) noch ausreicht, um die Behauptung (20) abzuleiten.

Im Spezialfall der Folgen ergibt sich aus Satz 3 eine Ergänzung zu Satz 2.

Satz 4. *Erfüllt eine Folge $\{a_n\}$ die Bedingung* (1) *und ist $a_n \geqq -K$ (bzw. $a_n \leqq K$) und $\overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|} \leqq 1$, so ist $s_n = \pm n + O(1)\,n^{\frac{1}{3}}(\log n)^{\frac{1}{3}}$ $(n\to\infty)$.*

Literatur

[1] Avakumović, V. G.: A note on a question set by P. Erdös and L. K. Hua. Publ. Inst. math. Beograd **6**, 47—56 (1954). — [2] Chandrasekharan, K., and S. Minakshisundaram: Typical means. Oxford 1952. — [3] Erdös, P.: Supplementary note. J. Ind. Math. Soc. **13**, 145—147 (1949). — [4] Widder, D. V.: The Laplace transform. Princeton 1946.

Trg. Republike 3/VI, Beograd (Jugoslavien)
Tübingen, Amselweg 18
Dept. of Math., Univ. of Cincinnati, Cincinnati, Ohio (USA)

(Eingegangen am 10. Oktober 1955)

DOLD, A.
Math. Zeitschr. Bd. 65, S. 200—206 (1956)

Vollständigkeit der Wuschen Relationen zwischen den Stiefel-Whitneyschen Zahlen differenzierbarer Mannigfaltigkeiten

Von

ALBRECHT DOLD

Einleitung. WU hat in [9] folgendes Verfahren zur Bestimmung der STIEFEL-WHITNEYschen Klassen W_i, $i=0,1,\dots,n$, einer n-dimensionalen kompakten differenzierbaren Mannigfaltigkeit M^n angegeben: Es bezeichne $H^*(M^n)=\sum_{\nu=0}^{n} H^\nu(M^n)$ den Kohomologiering mod 2 von M^n. Die Gruppen $H^p(M^n)$ und $H^{n-p}(M^n)$, $p=0,1,\dots,n$, sind bezüglich des Cup-Produktes $H^p(M^n)\times H^{n-p}(M^n)\to H^n(M^n)=Z_2$ (= zyklisch von der Ordnung 2) dual zueinander. Es gibt also eine eindeutig bestimmte Kohomologieklasse $U_p\in H^p(M^n)$ mit

$$U_p\cdot Y_{n-p}=Sq^p Y_{n-p}\ ^{1}) \quad \text{für alle } Y_{n-p}\in H^{n-p}(M^n). \tag{1}$$

Aus den U_p erhält man die W_i durch

$$W_i=\sum_{p=0}^{i} Sq^{i-p}U_p, \qquad i=0,1,\dots,n. \tag{2}$$

Diese WUsche Methode zeigt, daß zwischen den Klassen W_i Relationen bestehen. Es sei nämlich $P_q(W_i)$ ein homogenes Polynom in den W_i vom Gewicht q (das Gewicht eines Monoms $\prod_{i=0}^{n} W_i^{q_i}$ ist $\sum_{i=0}^{n} i q_i$). Nach (1) ist

$$U_{n-q}\cdot P_q(W_i)=Sq^{n-q}P_q(W_i). \tag{3}$$

Mit Hilfe der WUschen Formeln ([10], (1) oder [1], 7.1)

$$\begin{cases} Sq^r W_i=\sum_{t=0}^{r}\binom{i-r+t-1}{t} W_{r-t}\cdot W_{i+t},\ r\le i,\ W_\nu=0 \text{ für } \nu>n, \\ \binom{a}{b}=1 \text{ für } b=0,\quad \binom{a}{b}=0 \text{ für } a<b,\ b\neq 0 \end{cases} \tag{4}$$

und der CARTANschen Produktformel ([2], (5))

$$Sq^p(X\cdot Y)=\sum_{\nu+\mu=p} Sq^\nu X\cdot Sq^\mu Y \tag{5}$$

[1]) Mit Sq^p bezeichnen wir diejenige „square"-Kohomologieoperation mod 2 (s. [6] oder [7]), die die Dimension um p erhöht.

läßt sich die rechte Seite von (3) als Polynom (vom Gewicht n) in den W_i schreiben. Aus der [mit (2) identischen] Rekursionsformel $U_i = W_i + \sum_{p=0}^{i-1} Sq^{i-p} U_p$ erhält man wegen (4) und (5) eine Darstellung von U_p als Polynom in den W_i und damit die linke Seite von (3) als Polynom (vom Gewicht n) in den W_i. So ergibt sich z.B. für $P_q = W_{n-1}$ die Relation $W_1 \cdot W_{n-1} = W_1 \cdot W_{n-1} + n W_n$, d.h. $W_n = 0$ für ungerades n.

Die Relationen (3) zwischen den Klassen W_i sind von der besonderen Form $P_n(W_i) = 0$, wo $P_n(W_i)$ ein homogenes Polynom in den W_i vom Gewicht n ist. Die Werte, die die homogenen Polynome in den W_i vom Gewicht n auf dem Grundzykel der Mannigfaltigkeit M^n annehmen, nennt man die STIEFEL-WHITNEYschen Zahlen von M^n; (3) besagt also, daß gewisse STIEFEL-WHITNEYsche Zahlen in allen Mannigfaltigkeiten M^n Null sind. Ziel dieser Arbeit ist es, zu zeigen, daß jede Relation zwischen STIEFEL-WHITNEYschen Zahlen n-dimensionaler kompakter differenzierbarer Mannigfaltigkeiten Linearkombination von Relationen (3) ist.

Dieses Ergebnis stellt eine Teilantwort auf eine Frage von THOM ([8], S. 77–78) dar; der Beweis beruht auf einer Anwendung der Methoden und Ergebnisse von THOM [8].

I. Bezeichnungen, Vorbemerkungen. Es sei $\mathfrak{X}$ ein topologischer Raum. $H^*(\mathfrak{X}) = \sum_{\nu=0}^{\infty} H^\nu(\mathfrak{X})$ bezeichne den (singulären) Kohomologiering mod 2 von $\mathfrak{X}$. Wir setzen (zur Vereinfachung) voraus, daß $H^\nu(\mathfrak{X}) = 0$ ist für hinreichend großes ν. Ist $c \in H^*(\mathfrak{X})$, so bezeichnet c_ν die ν-dimensionale Komponente von c: $c_\nu \in H^\nu(\mathfrak{X})$, $c = \sum_{\nu=0}^{\infty} c_\nu$.

Ein Element c aus $H^*(\mathfrak{X})$ mit $c_0 = 1$ [= Einselement von $H^*(\mathfrak{X})$] besitzt ein eindeutig bestimmtes Inverses in $H^*(\mathfrak{X})$, nämlich $c^{-1} = \sum_{\mu=0}^{\infty} (1-c)^\mu$. (Beweis: $c \cdot \sum_{\mu=0}^{\infty} (1-c)^\mu = \big(1 - (1-c)\big) \sum_{\mu=0}^{\infty} (1-c)^\mu = 1$.)

Wir bezeichnen mit Sq die Summe der STEENRODschen Operationen $Sq^i : H^*(\mathfrak{X}) \to H^*(\mathfrak{X})$[1]) und bemerken:

$Sq = \sum_{i=0}^{\infty} Sq^i$ ist ein multiplikativer Automorphismus von $H^*(\mathfrak{X})$[2]).

Die CARTANsche Produktformel (5) besagt nämlich gerade, daß Sq ein multiplikativer Homomorphismus ist. Ist ferner $0 \neq c \in H^*(\mathfrak{X})$ und ν die kleinste ganze Zahl mit $c_\nu \neq 0$, so ist (wegen $Sq^0 =$ Identität) $(Sq\, c)_\nu = c_\nu \neq 0$, also $Sq\, c \neq 0$, also Sq ein Isomorphismus. Ist schließlich d ein beliebiges Element aus $H^*(\mathfrak{X})$ und $c = \sum_{\nu=0}^{\infty} c_\nu$ durch die Rekursionsformel

$$c_\nu = d_\nu + \Big(Sq \sum_{\varrho < \nu} c_\varrho\Big)_\nu \tag{6}$$

[2]) Diese für das Folgende sehr nützliche Bemerkung verdanke ich Herrn PUPPE; vgl. auch [3].

definiert, so ist $(Sqc)_\varrho = \left(Sq \sum_{\varrho \leq \varrho} c_\varrho\right)_\varrho = (Sqc_\varrho)_\varrho + \left(Sq \sum_{\varrho < \varrho} c_\varrho\right)_\varrho = c_\varrho + \left(Sq \sum_{\varrho < \varrho} c_\varrho\right)_\varrho = d_\varrho$, also $Sqc = d$, d.h. Sq ein Homomorphismus *auf* $H^*(\mathfrak{X})$.

II. Eigenschaften des Kohomologieringes mod 2 der GRASSMANNschen Mannigfaltigkeiten. Es bezeichne $O(l)$ die (reelle) orthogonale Gruppe in l Variablen und $G_{r,s}$ die GRASSMANNsche Mannigfaltigkeit $O(r+s)/O(r) \times O(s)$ (= Mannigfaltigkeit der nicht-orientierten r-dimensionalen Ebenen durch den Nullpunkt des euklidischen Raumes R^{r+s}); dabei sei $r \leq s$.

$G_{r,s}$ ist Basisraum einer $(s-1)$-universellen $O(r)$-Faserung

$$\mathfrak{B} = \{O(r+s)/O(s), p, G_{r,s}, O(r), O(r)\}$$

sowie einer $(r-1)$-universellen $O(s)$-Faserung

$$\overline{\mathfrak{B}} = \{O(r+s)/O(r), \bar{p}, G_{r,s}, O(s), O(s)\}$$

(s. [5], 19.7). w_i und $\overline{w}_j$, $i, j = 0, 1, \ldots$, seien die STIEFEL-WHITNEYschen Klassen mod 2 von $\mathfrak{B}$ und $\overline{\mathfrak{B}}$ ($w_0 = \overline{w}_0 = 1$, $w_i = 0$ für $i > r$, $\overline{w}_j = 0$ für $j > s$). Die Klassen w_i mit $1 \leq i \leq r$ sind bis zur Dimension $(s-1)$ algebraisch unabhängig und erzeugen $H^*(G_{r,s})$ (bis zur Dimension $s-1$), d.h. $H^*(G_{r,s})$ kann bis zur Dimension $(s-1)$ als Polynomalgebra in den Variablen w_i, $1 \leq i \leq r$, aufgefaßt werden; entsprechendes gilt für die $\overline{w}_j$ mit $1 \leq j \leq s$ (s. [1], Theorem 5.1). Wir benötigen den folgenden

Hilfssatz. *Sei* $w = \sum_{i=0}^{\infty} w_i$. *Durch die Festsetzungen*

$$w_i \to (w^{-1})_i, \quad i = 0, 1, \ldots \tag{7}$$

wird ein gradtreuer multiplikativer Automorphismus D von $H^*(G_{r,s})$ (bis zur Dimension $r-1$) definiert, der mit dem STEENRODschen *Automorphismus Sq* (s. I) *vertauschbar ist.*

Beweis. Wir beweisen zunächst die Gleichung $(w^{-1})_i = \overline{w}_i$, $i = 0, 1, \ldots$. Dazu bilden wir die $O(r+s)$-Hauptfaserung $\mathfrak{B} \oplus \overline{\mathfrak{B}}$ (= WHITNEYsches Produkt oder „join" von $\mathfrak{B}$ und $\overline{\mathfrak{B}}$). Sei $x \in G_{r,s}$. Die Faser in $\mathfrak{B}$ über x besteht aus allen in x gelegenen orthonormalen r-Beinen y; die in $\overline{\mathfrak{B}}$ aus allen auf x total orthogonalen, orthonormalen s-Beinen $\bar{y}$ des R^{r+s}. Die Faser in $\mathfrak{B} \oplus \overline{\mathfrak{B}}$ über x enthält die Paare $(y, \bar{y})$ und entsteht aus einem solchen Paar, indem man alle möglichen Rechtsverschiebungen (s. [5], 8.10) mit Elementen $g \in O(r+s)$ darauf anwendet. Faßt man $(y, \bar{y})$ als $(r+s)$-Bein, d.h. als Element von $O(r+s)$ auf, dann wird auch $(y, \bar{y}) \cdot g$ zu einem Element aus $O(r+s)$ und man erhält eine fasertreue Abbildung von $\mathfrak{B} \oplus \overline{\mathfrak{B}}$ in $O(r+s)$. Die Faserung $\mathfrak{B} \oplus \overline{\mathfrak{B}}$ ist also das Produkt ([5], § 10) und ihre STIEFEL-WHITNEYschen Klassen sind trivial. Nach dem WHITNEYschen Dualitätssatz (s. [11], Th. II) ist daher $\left(\sum_{i=0}^{\infty} w_i\right)\left(\sum_{j=0}^{\infty} \overline{w}_j\right) = 1$, d.h. $\overline{w}_i = (w^{-1})_i$.

Aus der oben beschriebenen Struktur von $H^*(G_{r,s})$ folgt nun, daß durch (7) tatsächlich ein gradtreuer Automorphismus D von $H^*(G_{r,s})$ (bis zur

Dimension $r-1$) definiert wird, und es bleibt nur noch die Vertauschbarkeit von D mit Sq nachzuweisen. Dafür genügt es zu zeigen, daß $D\,Sq(w_i) = Sq\,D(w_i)$ ist für $i = 0, 1, \ldots$. Nach den Formeln von Wu ([1], 7.1 oder (4)) ist $Sq(w_r)$ ein gewisses Polynom $P_{(i)}(w_r)$ in den Klassen w_r und ebenso ist $Sq(\overline{w}_r) = P_{(i)}(\overline{w}_r)$. Hieraus folgt $D\,Sq(w_i) = D\,P_{(i)}(w_r) = P_{(i)}(D(w_r)) = P_{(i)}(\overline{w}_r) = Sq(\overline{w}_i) = Sq\,D(w_i)$, q. e. d.

III. Die Wuschen Relationen. Es seien k, m, n natürliche Zahlen und $k > n$, $m \geq 2k$. Die Grassmannsche Mannigfaltigkeit $G_{k,m}$ (s. II) bezeichnen wir abgekürzt mit G_k; der spezielle Wert von m spielt im folgenden keine Rolle.

Ein Element aus $H^n(G_k)$, d.h. (s. II) ein homogenes Polynom $\pi_n(w_1, w_2, \ldots, w_n)$ in den Variablen w_i, $1 \leq i \leq n$, vom Gewicht n, bezeichnen wir als eine *Relation zwischen* Stiefel-Whitneyschen Zahlen und wir schreiben $\pi_n(w_1, w_2, \ldots, w_n) \equiv 0$, wenn in jeder n-dimensionalen kompakten differenzierbaren Mannigfaltigkeit M^n $\pi_n(W_1, W_2, \ldots, W_n) = 0$ ist (W_i die i-te Stiefel-Whitneysche Klasse von M^n).

Sei $w = \sum_{i=0}^{\infty} w_i$ und $u = Sq^{-1}w$ [s. I, (6)]. Dann ist nach Wu [s. (2) und (3)] $(u \cdot p + Sq\,p)_n \equiv 0$ für jedes $p \in H^*(G_k)$ [für $p \in H^j(G_k)$, $j \geq n$, ist sogar $(u \cdot p + Sq\,p)_n = 0$]. Wir behaupten nun umgekehrt

Satz 1. *Ein Element $\pi_n(w_1, w_2, \ldots, w_n)$ aus $H^n(G_k)$ ist genau dann eine Relation zwischen* Stiefel-Whitneyschen Zahlen, *wenn es sich in der Form $\pi_n(w_1, w_2, \ldots, w_n) = (u \cdot p + Sq\,p)_n$, $p \in H^*(G_k)$, schreiben läßt* $\left(Sq = \sum_{r=0}^{\infty} Sq^r, u = Sq^{-1}w = Sq^{-1}\sum_{i=0}^{\infty} w_i\right)$.

Setzen wir $p' = u \cdot p$, dann durchläuft mit p auch p' ganz $H^*(G_k)$ (weil u ein inversibles Element ist; s. I), und es ist $u \cdot p + Sq\,p = p' + Sq(u^{-1} \cdot p') = p' + (Sq\,u)^{-1} \cdot Sq\,p' = p' + w^{-1} \cdot Sq\,p'$, letzteres wegen $u = Sq^{-1}w$. Wir erhalten daher den zu Satz 1 äquivalenten

Satz 1'. *Ein Element $\pi_n(w_1, w_2, \ldots, w_n)$ aus $H^n(G_k)$ ist genau dann eine Relation zwischen* Stiefel-Whitneyschen Zahlen, *wenn es sich in der Form $\pi_n(w_1, w_2, \ldots, w_n) = (p + w^{-1} \cdot Sq\,p)_n$, $p \in H^*(G_k)$, schreiben läßt.*

Wir geben dem Satz 1 noch eine dritte, zu Satz 1' duale Formulierung, die wir dann im folgenden Abschnitt mit Hilfe der Ergebnisse von Thom [8] beweisen werden.

Jede n-dimensionale differenzierbare[3]) Mannigfaltigkeit M^n läßt sich differenzierbar in die $(n+k)$-Sphäre S^{n+k} einbetten (wegen $k > n$; s. z. B. [4], Th. 5, S. 12). Bezeichnet $W = \sum_{i=0}^{n} W_i$ die Stiefel-Whitneysche Klasse von M^n und $\overline{W} = \sum_{j=0}^{k} \overline{W}_j$ die Stiefel-Whitneysche Klasse des Faserraums der

[3]) Um uns den Thomschen Voraussetzungen in [8] anzupassen, benutzen wir von nun an das Wort „differenzierbar" im Sinne von „unendlich oft differenzierbar"; es genügt offenbar, die Vollständigkeit der Wuschen Relationen für unendlich oft differenzierbare Mannigfaltigkeiten zu beweisen.

Normalenvektoren von M^n in S^{n+k} (bei einer solchen Einbettung), dann ist nach der WHITNEYschen Dualitätsformel ([11], Th. II) $W \cdot \overline{W} = 1$ (denn die STIEFEL-WHITNEYsche Klasse von S^{n+k} ist 1), d.h. $\overline{W} = W^{-1}$. Ist nun $\pi_n(w_1, w_2, \ldots, w_n) \in H^n(G_k)$, so folgt hieraus

$$\pi_n(W_1, W_2, \ldots, W_n) = \pi_n((\overline{W}^{-1})_1, (\overline{W}^{-1})_2, \ldots, (\overline{W}^{-1})_n) = D\,\pi_n(\overline{W}_1, \overline{W}_2, \ldots, \overline{W}_n)$$

(Definition des Automorphismus D im Hilfssatz, Abschnitt II). Insbesondere ist $\pi_n(w_1, w_2, \ldots, w_n)$ genau dann eine Relation zwischen STIEFEL-WHITNEYschen Zahlen, wenn in jeder n-dimensionalen kompakten differenzierbaren Untermannigfaltigkeit von S^{n+k} $D\pi_n(\overline{W}_1, \overline{W}_2, \ldots, \overline{W}_n) = 0$ ist. Bezeichnen wir ein homogenes Polynom $\bar{\pi}_n(w_1, w_2, \ldots, w_n)$ vom Gewicht n mit der letztgenannten Eigenschaft als *„Relation zwischen normalen* STIEFEL-WHITNEY*schen Zahlen"*, so führt also der Automorphismus D die Relationen zwischen (tangentialen) STIEFEL-WHITNEYschen Zahlen gerade in die Relationen zwischen normalen STIEFEL-WHITNEYschen Zahlen über. Andererseits führt er die Elemente der Form $(p + w^{-1} \cdot Sqp)_n$, $p \in H^*(G_k)$, in $D((p + w^{-1} \cdot Sqp)_n) = (D(p + w^{-1} \cdot Sqp))_n = (Dp + Dw^{-1} \cdot DSqp)_n = (Dp + w \cdot SqDp)_n$ über, und Dp durchläuft mit p ganz $H^*(G_k)$ (bis zur Dimension $k-1$). Somit ergibt sich der zu Satz 1′ äquivalente

Satz 1″. *Ein Element $\bar{\pi}_n(w_1, w_2, \ldots, w_n)$ aus $H^n(G_k)$ ist genau dann eine Relation zwischen normalen* STIEFEL-WHITNEY*schen Zahlen, wenn es sich in der Form*

$$\bar{\pi}_n(w_1, w_2, \ldots, w_n) = (p + w \cdot Sqp)_n, \quad p \in H^*(G_k), \tag{8}$$

schreiben läßt $\left(w = \sum_{i=0}^{\infty} w_i,\ Sq = \sum_{r=0}^{\infty} Sq^r\right)$.

IV. Beweis des Satzes 1″. Wir stellen zunächst einige Ergebnisse von THOM [8] zusammen, die wir zum Beweis des Satzes 1″ benötigen. Die Bezeichnungen und Voraussetzungen sind dieselben wie im Abschnitt III.

THOM konstruiert mit Hilfe des $(2k-1)$-universellen $(k-1)$-Sphärenfaserraumes über G_k

1. eine Pseudomannigfaltigkeit $MO(k)$, die G_k enthält (s. [8], II. 2),

2. einen (additiven) Isomorphismus φ^* von $H^*(G_k)$ auf $H^*_+(MO(k))$ [= Untergruppe der Elemente positiver Dimension von $H^*(MO(k))$], der die Dimension homogener Elemente um k erhöht (s. [8], II. 2),

3. zu jeder n-dimensionalen kompakten differenzierbaren Untermannigfaltigkeit $M \subset S^{n+k}$ eine bis auf Homotopie bestimmte stetige Abbildung $F_M: S^{n+k} \to MO(k)$, die M in G_k überführt; $f_M: M \to G_k$ bezeichne die durch F_M definierte Abbildung (s. [8], II. 3 oder IV. 4; in IV. 4 hat man von der Orientierung der Mannigfaltigkeiten abzusehen, $MSO(k)$ durch $MO(k)$ und $\tilde{G}_k$ durch G_k zu ersetzen).

Es gilt:

A. α) Der durch f_M induzierte Homomorphismus $f_M^*: H^*(G_k) \to H^*(M)$ führt w_i, $i = 0, 1, \ldots, k$, in die i-te STIEFEL-WHITNEYsche Klasse $\overline{W}_i$ des

Normalenraumes von M in S^{n+k} über. (Aus der Definition von F_M in [8], IV. 4 geht unmittelbar die schärfere Aussage hervor: Der durch f_M induzierte $(k-1)$-Sphärenfaserraum über M ist mit dem Raum der normalen Einheitsvektoren von M äquivalent.)

β) Das Diagramm

$$\begin{array}{ccc} H^*(MO(k)) & \xrightarrow{F_M^*} & H^*(S^{n+k}) \\ \varphi^* \uparrow & & \uparrow \psi^* \\ H^*(G_k) & \xrightarrow[f_M^*]{} & H^*(M) \end{array} \tag{9}$$

ist kommutativ. Dabei ist F_M^* der durch F_M induzierte Homomorphismus, und ψ^* ist der zur Injektion $M \to S^{n+k}$ gehörige Umkehrhomomorphismus; ψ^* führt die Grundklasse $m \in H^n(M)$ in die Grundklasse $s \in H^{n+k}(S^{n+k})$ über (s. [8], II. 3; dort sind im Diagramm (1) die Richtungen der horizontalen Pfeile umzukehren).

B. Der durch die Injektion induzierte Homomorphismus $i^*: H^*(MO(k)) \to H^*(G_k)$ bildet (in den Dimensionen >0 und $<2k$) $H_+^*(MO(k))$ isomorph auf das von w_k in $H^*(G_k)$ erzeugte Ideal ab. Wir identifizieren daher $H_+^*(MO(k))$ vermöge i^* mit diesem Ideal. Der Homomorphismus φ^* führt $p \in H^*(G_k)$, $\dim(p) < k$, in $p \cdot w_k$ über ([8], II. 5).

C. $H^{n+k}(MO(k))$ besitzt eine Basis, die sich aus zwei Teilsystemen $\{a^{(i)}\}$ und $\{b^{(j)}\}$ mit den folgenden Eigenschaften zusammensetzt (d.i. die Basis (7) in [8], II. 6; die $a^{(i)}$ sind die Elemente X_ω^{n+k}, die $b^{(j)}$ die übrigen):

α) Ist a eine nicht-triviale Linearkombination von Elementen $a^{(i)}$, so gibt es eine Mannigfaltigkeit $M \subset S^{n+k}$ mit (vgl. A. β) $F_M^*(a) = s \neq 0$ ([8], IV. 6).

β) $b^{(j)} = Sq^{r_j} p^{(j)}$ für geeignetes $p^{(j)} \in H^*(MO(k))$ und $r_j > 0$.

Wir kommen nun zum eigentlichen Beweis des Satzes 1''. Es sei S_k^n die Untergruppe der Elemente c von $H^{n+k}(MO(k))$ mit der Eigenschaft: Bei jeder Abbildung F_M (s. 3.) ist $F_M^*(c) = 0$ (s. A. β). Da der Automorphismus Sq in $H^*(S^{n+k})$ die Identität ist und weil Sq mit F_M^* vertauschbar ist, liegen alle Elemente der Form $(Sq\,p' + p')_{n+k}$, $p' \in H^*(MO(k))$, in S_k^n, insbesondere also die Elemente $b^{(j)}$ aus C. Nach C. α) liegt keine nicht-triviale Linearkombination der $a^{(i)}$ in S_k^n, d.h. S_k^n wird von den $b^{(j)}$ erzeugt und besteht genau aus den Elementen der Form $(Sq\,p' + p')_{n+k}$, $p' \in H^*(MO(k))$.

Betrachten wir nun das Diagramm (9). $\varphi^*: H^n(G_k) \to H^{n+k}(MO(k))$ und $\psi^*: H^n(M) \to H^{n+k}(S^{n+k})$ sind Isomorphismen. Aus der Kommutativität von (9) folgt daher: $\varphi^{*-1}(S_k^n)$ ist die Untergruppe der Elemente $\bar{\pi}$ aus $H^n(G_k)$ mit der Eigenschaft: Bei jeder Abbildung f_M (s. 3.) ist $f_M^*(\bar{\pi}) = 0$. Wegen $f_M^*(w_i) = \overline{W}_i$ (s. A. α) besteht $\varphi^{*-1}(S_k^n)$ gerade aus den Relationen zwischen normalen STIEFEL-WHITNEYschen Zahlen. Wir erhalten also: Ein Element $\bar{\pi}(w_1, w_2, \ldots, w_k)$ aus $H^n(G_k)$ ist genau dann eine Relation zwischen normalen STIEFEL-WHITNEYschen Zahlen, wenn es sich in der Form

$$\varphi^{*-1}\big((Sq\,p' + p')_{n+k}\big), \quad p' \in H^*(MO(k)), \tag{10}$$

schreiben läßt. Zum Beweis des Satzes 1'' haben wir jetzt lediglich noch (10) in (8) umzurechnen.

$H_+^*(MO(k))$ besteht (bis zur Dimension $2k-1 \geq n+k$) aus allen Elementen der Form $p' = p \cdot w_k$, $p \in H^*(G_k)$ (s. B.). Aus den Wuschen Formeln [vgl. (4)]

$$Sq^r w_i = \sum_{t=0}^{r} \binom{i-r+t-1}{t} w_{r-t} \cdot w_{i+t} \text{ folgt } Sq^r w_k = w_r \cdot w_k \ (w_r = 0 \text{ für } r > k!),$$

d.h. $Sq w_k = w \cdot w_k$. Daher ist

$$Sq p' + p' = Sq(p \cdot w_k) + p \cdot w_k = Sq p \cdot Sq w_k + p \cdot w_k = (w \cdot Sq p + p) \cdot w_k,$$

also $\varphi^{*-1}(Sq p' + p') = w \cdot Sq p + p$ (s. B.). φ^* erhöht die Dimension um k (s. 2.), also ist $\varphi^{*-1}((Sq p' + p')_{n+k}) = (\varphi^{*-1}(Sq p' + p'))_n = (w \cdot Sq p + p)_n$, q.e.d.

Literatur

[1] BOREL, A.: La cohomologie de certains espaces homogènes. Comment. Math. Helv. **27**, 165—197 (1953). — [2] CARTAN, H.: Une théorie axiomatique des carrés de Steenrod. C. R. Acad. Sci. Paris **230**, 425—427 (1950). — [3] PUPPE, D.: Über Homotopiemengen und ihre induzierten Abbildungen. (Erscheint demnächst.) — [4] RHAM, G. DE: Variétés différentiables. Act. sci. industr. 1222, Paris. — [5] STEENROD, N.: The topology of fibre bundles. Princeton Math. Ser. 14. — [6] STEENROD, N.: Products of cocycles and extensions of mappings. Ann. of Math. **48**, 290—320 (1947). — [7] STEENROD, N.: Homology groups of symmetric groups and reduced power operations; cyclic reduced powers of cohomology classes. Proc. Nat. Acad. USA. **39**, 217—223 (1953). — [8] THOM, R.: Quelques propriétés globales des variétés différentiables. Comment. Math. Helv. **28**, 17—86 (1954). — [9] WU, WEN-TSÜN: Classes caractéristiques et i-carrés d'une variété. C. R. Acad. Sci. Paris **230**, 508—511 (1950). — [10] WU, WEN-TSÜN: Les i-carrés dans une variété grassmannienne. C. R. Acad. Sci. Paris **230**, 918—920 (1950). — [11] WU, WEN-TSÜN: On the product of sphere-bundles and the duality theorem modulo two. Ann. of Math. **49**, 641—653 (1948).

Heidelberg, Mathematisches Institut der Universität

(Eingegangen am 22. November 1955)

MOHANTY, R., and S. MOHAPATRA
Math. Zeitschr. Bd. 65, S. 207—213 (1956)

On the absolute logarithmic summability of a Fourier series

By

R. MOHANTY and S. MOHAPATRA

1. Let $f(t)$ be L integrable in $(-\pi, \pi)$ and periodic with period 2π and let

$$f(t) \sim \tfrac{1}{2}a_0 + \sum_1^\infty (a_n \cos nt + b_n \sin nt) = \tfrac{1}{2}a_0 + \sum_1^\infty A_n(t). \tag{1.1}$$

We write

$$\varphi(t) = \tfrac{1}{2}\{f(x+t) + f(x-t) - 2f(x)\}, \tag{1.2}$$

$$\Phi(t) = \int_t^\pi \frac{\varphi(u)}{u}\,du = \varphi_1(t)\log\frac{2\pi}{t}. \tag{1.3}$$

DEFINITION. Let $\lambda = \lambda(w)$ be continuous, differentiable and monotonic increasing in $(0, \infty)$ and let it tend to infinity as $w \to \infty$. A series $\sum a_n$ is summable $|R, \lambda, r|$ where $r > 0$, if $\int_A^\infty \frac{r\lambda'(w)}{\{\lambda(w)\}^{r+1}} \left|\sum_{n<w} \{\lambda(w) - \lambda(n)\}^{r-1} a_n\right| < \infty$ where A is a positive fixed number [2].

The object of the present paper is to prove the following

THEOREM. *If $\varphi_1(t)$ is of bounded variation in $(0, \pi)$, then (1.1) is summable $|R, \log w, k|$ for every $k > 1$.*

REMARK. The order of summability in the conclusion of the above theorem can not be replaced by 1 as $|R, \log w, 1|$ summability of a Fourier series is a non-local property of its generating function [3] and therefore can not be ensured by the hypothesis of the theorem, which is obviously a local one.

2. In the proof of the theorem we require the following inequalities for the kernel $h(w, t) = \sum_{n<w} \log n \left(\log\frac{w}{n}\right)^{k-1} \cos nt$.

$$h(w, t) = O\{t^{-1} \log w (\log wt)^{k-2}\}, \tag{2.1}$$

$$h(w, t) = O\{w(\log w)^k\}. \tag{2.2}$$

The proof of (2.2) is trivial. For the proof of (2.1) we require the following lemmas.

LEMMA 1. $\int_\alpha^\beta \sin\left(\left[\frac{u}{t}\right] + \frac{1}{2}\right)t\,du = O(1)$ *for* $0 < t < \pi$, $0 < \alpha \leq t$; *and* $\alpha < \beta \leq wt$.

PROOF. We have

$$\int_\alpha^\beta \sin\left(\left[\frac{u}{t}\right]+\frac{1}{2}\right)t\,du = \int_\alpha^{t\left[\frac{\alpha}{t}\right]+t} + \int_{t\left[\frac{\alpha}{t}\right]+t}^{t\left[\frac{\beta}{t}\right]} + \int_{t\left[\frac{\beta}{t}\right]}^{\beta}$$

$$= O(1) + t \sum_{\left[\frac{\alpha}{t}\right]+1 \le n < \left[\frac{\beta}{t}\right]} \sin\left(n+\tfrac{1}{2}\right)t + O(1)$$

$$= O(1) + t\,O(t^{-1}) + O(1) = O(1),$$

for $0<t<\pi$ *and* $t \le \alpha < \beta \le wt$.

LEMMA 2.

$$\int_\alpha^\beta \sin\left(\left[\frac{u}{t}\right]+\frac{1}{2}\right)t\,\frac{du}{u} = O(1)$$

for $0<t<\pi$ and $t \le \alpha < \beta \le wt$.

PROOF.

$$\int_\alpha^\beta = \int_\alpha^\pi + \int_\pi^\beta = P + Q.$$

By the second mean value theorem and Lemma 1, we have

$$Q = \frac{1}{\pi}\int_\pi^{\beta'} \sin\left(\left[\frac{u}{t}\right]+\frac{1}{2}\right)t\,du = O(1)$$

where $\pi<\beta'<\beta$.

Now

$$|P| = \left|\int_\alpha^\pi \sin\left(\left[\frac{u}{t}\right]+\frac{1}{2}\right)t\,\frac{du}{u}\right| \le O(1)\int_t^\pi \left(u+O(t)\right)\frac{du}{u}$$

$$= O\{(\pi - t)\} + O(t\log t) = O(1).$$

Proof of (2.1). We have

$$\sum_{1\le n<w} \left(\log\frac{w}{n}\right)^{k-1} \log n \cos nt = \int_1^w \frac{d}{dx}\left\{\left(\log\frac{w}{x}\right)^{k-1}\log x\right\}\left(\frac{1}{2}+\sum_{1\le n<x}\cos nt\right)$$

$$= \frac{1}{2\sin\frac{1}{2}t}\int_1^w \sin\left([x]+\frac{1}{2}\right)t\left\{\frac{\left(\log\frac{w}{x}\right)^{k-1}}{x} - (k-1)\frac{\left(\log\frac{w}{x}\right)^{k-2}}{x}\log x\right\}dx$$

$$= \frac{1}{2\sin\frac{1}{2}t}\left[k\int_1^w \frac{\sin\left([x]+\frac{1}{2}\right)t}{x}\left(\log\frac{w}{x}\right)^{k-1}dx - \right.$$

$$\left. - (k-1)\log w\int_1^w \frac{\left(\log\frac{w}{x}\right)^{k-2}}{x}\sin\left([x]+\frac{1}{2}\right)t\,dx\right]$$

$$= \frac{1}{2\sin\frac{1}{2}t}\left[k\int_t^{wt} \sin\left(\left[\frac{u}{t}\right]+\frac{1}{2}\right)t\left(\log\frac{wt}{u}\right)^{k-1}\frac{du}{u} - \right.$$

$$\left. - (k-1)\log w\int_t^{wt} \sin\left(\left[\frac{u}{t}\right]+\frac{1}{2}\right)t\left(\log\frac{wt}{u}\right)^{k-2}\frac{du}{u}\right]$$

$$= \frac{1}{2\sin\frac{1}{2}t}\left[I_1 - \log w\, I_2\right].$$

Now

$$I_1 = k\int_t^{wt}\left(\log\frac{wt}{u}\right)^{k-1}\sin\left(\left[\frac{u}{t}\right]+\frac{1}{2}\right)t\,\frac{du}{u}$$

$$= k\,(\log w)^{k-1}\int_t^{\beta}\sin\left(\left[\frac{u}{t}\right]+\frac{1}{2}\right)t\,\frac{du}{u},$$

and hence by the second mean value theorem and Lemma 2,

$$I_1 = O\{(\log w)^{k-1}\}. \tag{2.3}$$

Again

$$I_2 = \int_t^{wt} = \int_t^{wt-1} + \int_{wt-1}^{wt} = P_1 + Q_1.$$

$$P_1 = \int_t^{wt-1} = \int_t^{\mu} + \int_{\mu}^{\eta wt} + \int_{\eta wt}^{wt-1} = P_2 + P_3 + P_4,$$

where μ is a constant and η has been so chosen that $\dfrac{\left(\log\frac{wt}{u}\right)^{k-2}}{u}$ is monotonic decreasing in $0<u<\eta\,wt$ and also monotonic increasing in $\eta\,wt\leq w<wt-1$, η lying between 0 and 1.

Therefore by the second mean value theorem, since $\left(\log\frac{wt}{u}\right)^{k-2}$ is monotonic increasing

$$\left\{\begin{aligned} P_2 &= \left(\log\frac{wt}{\mu}\right)^{k-2}\int_t^{\mu}\sin\left(\left[\frac{u}{t}\right]+\frac{1}{2}\right)t\,\frac{du}{u}\\ &= O\{(\log wt)^{k-2}\}, \quad \text{by Lemma 2.}\end{aligned}\right. \tag{2.4}$$

Applying the second mean value theorem again,

$$\left\{\begin{aligned} P_3 &= \int_{\mu}^{\eta wt}\frac{\left(\log\frac{wt}{u}\right)^{k-2}}{u}\sin\left(\left[\frac{u}{t}\right]+\frac{1}{2}\right)t\,du\\ &= \frac{\left(\log\frac{wt}{\mu}\right)^{k-2}}{\mu}\int_{\mu}^{\xi}\sin\left(\left[\frac{u}{t}\right]+\frac{1}{2}\right)t\,du \qquad (\mu<\xi<\eta\,wt)\\ &= O\{(\log wt)^{k-2}\} \quad \text{by Lemma 1.}\end{aligned}\right. \tag{2.5}$$

Since $\dfrac{\left(\log\frac{wt}{u}\right)^{k-2}}{u}$ is monotonic increasing in $\eta\,wt<u<wt-1$ we have by the second mean value theorem and Lemma 1,

$$P_4 = O\left\{\frac{\left(\log\frac{wt}{wt-1}\right)^{k-2}}{wt-1}\right\} = O\left\{\frac{1}{(wt-1)^{k-1}}\right\} = O\left\{\frac{1}{(wt)^{k-1}}\right\}. \tag{2.6}$$

Thus from (2.4), (2.5), and (2.6),

$$P_1 = O\{(\log wt)^{k-2}\}. \tag{2.7}$$

Now

$$|Q_1| = \left| \int_{wt-1}^{wt} \frac{\left(\log \frac{wt}{u}\right)^{k-2}}{u} \sin\left(\left[\frac{u}{t}\right] + \frac{1}{2}\right) t\, du \right|$$

$$\leq \int_{wt-1}^{wt} \frac{\left(\log \frac{wt}{u}\right)^{k-2}}{u}\, du$$

$$= O\left\{\left(\log \frac{wt}{wt-1}\right)^{k-1}\right\} = O\left\{\frac{1}{(wt)^{k-1}}\right\}.$$

Thus

$$I_2 = O\{(\log wt)^{k-2}\} \tag{2.8}$$

and from (2.3) and (2.8)

$$h(w,t) = O\{t^{-1} \log w (\log wt)^{k-2}\}.$$

3. PROOF OF THE THEOREM. On integrating by parts twice in succession we find

$$\left\{\begin{aligned} A_n(x) &= \frac{2}{\pi}\int_0^{\pi} \varphi(t) \cos nt\, dt = -\frac{2}{\pi}\int_0^{\pi} t\, \Phi'(t) \cos nt\, dt \\ &= -\frac{2}{\pi}\int_0^{\pi} d\{\varphi_1(t)\}\left\{t \log \frac{2\pi}{t} \cos nt + \frac{\sin nt}{n}\right\}, \\ &= u_n + v_n. \end{aligned}\right. \tag{3.1}$$

The series $\sum u_n$ is summable $|R, \log w, k|$ if

$$I = \frac{2}{\pi}\int_0^{\pi} \frac{k\, dw}{w(\log w)^{k+1}} \left| \sum_{n<w} \log n \left(\log \frac{w}{n}\right)^{k-1} u_n \right| < \infty$$

substituting for u_n from (3.1) it is easy to see that

$$I \leq \frac{2k}{\pi}\int_0^{\pi} J\, |d\varphi_1(t)|\, t \log \frac{2\pi}{t}, \tag{3.2}$$

where

$$J = \int_A^{\infty} \frac{dw}{w(\log w)^{k+1}} |h(w,t)|.$$

Since $\int_0^{\pi} |d\varphi_1(t)|$ is finite, it is enough to show that

$$J = O\left\{t^{-1}\left(\log \frac{2\pi}{t}\right)^{-1}\right\} \quad \text{for} \quad 0 < t < \pi.$$

Now

$$\left\{\begin{aligned} J &= \int_A^{\infty} \frac{dw}{w(\log w)^{k+1}} |h(w,t)| \\ &= \int_A^{2\pi/t} + \int_{2\pi/t}^{\infty} = J_1 + J_2. \end{aligned}\right. \tag{3.3}$$

Employing (2.2), we have

$$J_1 < C \int_A^{2\pi/t} \frac{dw}{\log w} = O\left\{t^{-1}\left(\log \frac{2\pi}{t}\right)^{-1}\right\}. \tag{3.4}$$

Employing (2.1), we have

$$J_2 < C\, t^{-1} \int_{2\pi/t}^{\infty} \frac{(\log w\, t)^{k-2}}{w(\log w)^k}\, dw, \tag{3.5}$$

and

$$\left\{\begin{aligned} \int_{2\pi/t}^{\infty} \frac{(\log w\, t)^{k-2}}{w(\log w)^k}\, dw &= \left[-\frac{\left(1 - \dfrac{\log \frac{1}{t}}{\log w}\right)^{k-1}}{(k-1)\log \frac{1}{t}} \right]_{2\pi/t}^{\infty} \\ &= O\left\{\left(\log \frac{2\pi}{t}\right)^{-1}\right\}. \end{aligned}\right. \tag{3.6}$$

Hence from (3.5) and (3.6)

$$J_2 = O\left\{t^{-1}\left(\log \frac{2\pi}{t}\right)^{-1}\right\},$$

and therefore by (3.4) and (3.6), we have

$$J = O\left\{t^{-1}\left(\log \frac{2\pi}{t}\right)^{-1}\right\}. \tag{3.7}$$

Now

$$v_n = \frac{2}{\pi} \int_0^{\pi} d\{\varphi_1(t)\} \frac{\sin nt}{n}.$$

Since $\varphi_1(t)$ is of bounded variation, $\sum v_n$ is summable $|C, \delta|$, $\delta > 0$, [1] and hence it is summable $|C, 1+\delta|$ and a fortiori summable $|R, \log w, 1+\delta|$, $\delta > 0$ i.e. summable $|R, \log w, k|$, $k > 1$.

Thus the proof of the theorem is complete.

4. The integral I_2 in section 2 is analogous to the integral

$$\int_0^{\lambda} \left(\log \frac{\lambda}{u}\right)^k \frac{\sin u}{u}\, du, \qquad -1 < k < 0,$$

which naturally suggests the problem of $(R, \log w, k)$ summability of $\int_0^\infty \frac{\sin u}{u}\,du$, where $-1<k<0$. That the integral is bounded $(R, \log w, k)$ is implied in a result due to WANG [4].

For proving the summability $(R, \log w, k)$, $-1<k<0$, of $\int_0^\infty \frac{\sin u}{u}\,du$ to the value $\frac{\pi}{2}$, we have only to show that

$$\int_0^\lambda \left(\log \frac{\lambda}{u}\right)^k \frac{\sin u}{u}\,du \sim \frac{\pi}{2}(\log \lambda)^k. \tag{4.1}$$

The proof of this is very similar to that of the result due to WANG referred to above and is sketched below for the sake of completeness.

We write

$$\int_0^\lambda \left(\log \frac{\lambda}{u}\right)^k \frac{\sin u}{u}\,du = \int_0^\mu + \int_\mu^{\lambda e^k} + \int_{\lambda e^k}^{\lambda e^{-\frac{1}{\lambda}}} + \int_{\lambda e^{-\frac{1}{\lambda}}}^{\lambda}$$

$$= P_1 + P_2 + P_3 + P_4.$$

Now

$$P_1 = \int_0^\mu \left(\log \frac{\lambda}{u}\right)^k \frac{\sin u}{u}\,du = (\log \lambda)^k \int_0^\mu \left(1 - \frac{\log u}{\log \lambda}\right)^k \frac{\sin u}{u}\,du$$

$$= (\log \lambda)^k P_1^*.$$

Since

$$P_1^* \to \int_0^\mu \frac{\sin u}{u}\,du = \frac{\pi}{2}\left\{1 + O\left(\frac{1}{\mu}\right)\right\}$$

$$P_1 \sim \frac{\pi}{2}(\log \lambda)^k. \tag{4.2}$$

$$P_2 = \frac{\left(\log \frac{\lambda}{\mu}\right)^k}{\mu} \int_\mu^\xi \sin u\,du = O\left\{\frac{(\log \lambda)^k}{\mu}\right\}, \tag{4.3}$$

where $\mu<\xi<\lambda e^k$, since $\frac{\left(\log \frac{\lambda}{u}\right)^k}{u}$ is monotonic decreasing.

$$P_3 = \left(\log e^{-\frac{1}{\lambda}}\right)^k \frac{1}{\lambda} e^{\frac{1}{\lambda}} \int_{\xi'}^{\lambda e^{-\frac{1}{\lambda}}} \sin u\,du = O\left(\frac{1}{\lambda^{k+1}}\right), \quad \left(\lambda e^k < \xi' < \lambda e^{-\frac{1}{\lambda}}\right), \tag{4.4}$$

$$|P_4| \leq \int_{\lambda e^{-\frac{1}{\lambda}}}^{\lambda} \frac{\left(\log \frac{\lambda}{u}\right)^k}{u}\,du = \left(\log e^{\frac{1}{\lambda}}\right)^{k+1} = O\left(\frac{1}{\lambda^{k+1}}\right) \tag{4.5}$$

and therefore by (4.2), (4.3), (4.4) and (4.5)

$$\int_0^{\lambda} \left(\log \frac{\lambda}{u}\right)^k \frac{\sin u}{u}\, du \sim \frac{\pi}{2} (\log \lambda)^k.$$

We are thankful to the referee for his kind suggestions in obtaining the requisite inequality for the kernel $h(w, t)$ and for certain other remarks

References

[1] Bosanquet, L. S.: Note on the absolute summability (C) of a Fourier series. J. London Math. Soc. (1) **11**, 11—15 (1936). — [2] Mohanty, R.: On the absolute Riesz summability of a Fourier series and its allied series. Proc. London Math. Soc. (2) **52**, 295—320 (1951). — [3] Mohanty, R.: On the summability $|R, \log w, 1|$ of a Fourier series. J. London Math. Soc. **25**, 67—72 (1950). — [4] Wang, F. T.: On the summability of Fourier series by Riesz's Logarithmic mean, II. Tohoku Math. J. **40**, 273—292 (1935).

Department of Mathematics, Ravenshaw College, Cuttack (India)

(Eingegangen am 23. November 1955)

EICHLER, M.
Math. Zeitschr. Bd. 65, S. 214 (1956)

Berichtigung zu meiner Arbeit über den Hilbertschen Klassenkörper eines imaginärquadratischen Zahlkörpers

Von
MARTIN EICHLER

Durch unglückliche Umstände wurden meine Korrekturen zu der oben genannten, in Band 64, S. 229—242 erschienenen Arbeit nicht ausgeführt. Sie müssen nun nachträglich als Berichtigung erscheinen.

S. 229. Die erste Formel lautet

$$F_n(x, x') = \prod\left(x - j\left(\frac{r_1\tau' + r_2}{r_3}\right)\right).$$

S. 231. Z. 16, 17 von unten. Der hier stehende Satz wird in folgender Form verständlicher: „Das der (wie gezeigt eindeutig fixierten) Isomorphie $\mathfrak{o} \cong \mathfrak{o}_i$ dem $\mathfrak{o}$-Ideal $\mathfrak{a}$ entsprechende $\mathfrak{o}_i$-Ideal heiße $\mathfrak{a}_i$." — Die im folgenden auftretenden Permutationen sollen mit σ'_C, σ'_{C_1} usw. (also mit einem Akzent) bezeichnet werden.

S. 233. Z. 5 an Stelle von „wie gezeigt" lies „wie vorausgesetzt wird". — Der zweite Absatz auf dieser Seite ist durch folgenden Text zu ersetzen: „Wenn $\mathfrak{o}_0$ im Gegensatz zur Behauptung nicht mit $\mathfrak{o}$ isomorph wäre, so wäre $\mathfrak{o}_0$ entweder mit einer echten Teilordnung von $\mathfrak{o}$ isomorph ($\mathfrak{o}$ ist maximal!), oder aber die von $\mathfrak{o}_0$ und $\mathfrak{o}$ erzeugten quadratischen Zahlkörper wären nicht isomorph. Beides führt auf einen Widerspruch. Im letzteren Falle ist dieses fast evident. Im ersteren Falle nehmen wir ohne Beschränkung der Allgemeinheit $\mathfrak{o}_0 < \mathfrak{o}$ an. Es seien $[1, \omega]$ mit $s(\omega) = 0$ oder 1 und $[1, f\omega]$ Basen von $\mathfrak{o}$ und $\mathfrak{o}_0$; dabei ist f eine natürliche Zahl > 1. Die kleinsten Normen von nicht rationellen Elementen aus $\mathfrak{o}$ und $\mathfrak{o}_0$ sind $n(\omega)$ und $f^2 n(\omega)$. Mithin gibt es kein Element in $\mathfrak{o}_0$ der Norm $n(\omega)$, im Widerspruch zur Voraussetzung."

S. 235. Z. 18—19 ersetze durch:

$$\xi_1^{\sigma'_C} \equiv \xi_1^{p^{-e}} \bmod \mathfrak{P}_1.$$

Dies gilt auch unter Vertauschung von ξ_1 und $\xi_1^{\sigma'_C}$. Also kann man — Z. 9 von unten lies $\xi_3^{\sigma'_C} \equiv \xi_3^p \bmod \mathfrak{P}_3$. — Z. 8 von unten lies σ'_C an Stelle von σ_C.

S. 236. Z. 9 lies $(\xi_i^{\sigma'_C} - \xi_i^{p^2})(\xi_i^{p^2\sigma'_C} - \xi_i)$. Ferner Z. 5 von unten: Im gleichen Zuge ergibt sich, zunächst für $f > 1$: Hilfssatz 2. *Für eine Klasse C ist $\sigma_C^{\pm 1}$ der Frobenius-Automorphismus eines Primteilers $\mathfrak{P}$ von p in K, wobei p die Norm eines Primideals $\mathfrak{p}$ ersten Grades aus der Klasse C ist. Das gilt für fast alle $\mathfrak{p}$ dieser Klasse. Usw.*"

S. 241. Z. 5 lies „in k" an Stelle von „von K". — Z. 7 lies „Hilfssatz 6" statt „Hilfssatz 2". — Z. 7 von unten lies „Hilfssatz 6" statt „Hilfssatz 5". — Z. 9 von unten bis zum Schluß: lies σ_C an Stelle von σ'_C und σ_0 an Stelle von σ'_0. — Letzte Z. lies σ'_C an Stelle von σ_C.

S. 242. Z. 1 und Z. 7 lies σ_C und σ_0 an Stelle von σ'_C und σ'_0. — Z. 12 lautet richtig $\xi_i^{\sigma_C} \equiv \xi_i^p \bmod \mathfrak{p}$, $\xi_i^{\sigma_C} \equiv \xi_i^p \bmod \bar{\mathfrak{p}}$.

Münster i. Westf., Rosenstr. 9

HUCKEMANN, F.
Math. Zeitschr. Bd. 65, S. 215—239 (1956)

Zur Darstellung von Riemannschen Flächen durch Streckenkomplexe*)

Von
FRIEDRICH HUCKEMANN

Einleitung

§ 1. Bekannt ist die Konstruktion des Streckenkomplexes für schlichtartige (einfach zusammenhängende) RIEMANNsche Flächen, deren Windungspunkte sich in endlich viele Grundpunkte projizieren[1]): Man verbindet diese Punkte durch eine geschlossene doppelpunktfreie JORDAN-Kurve L und zerschneidet die Fläche über L in allen ihren Blättern; sie zerfällt so in (im allgemeinen unendlich viele) Halbblätter, *Innengebiete* und *Außengebiete*, deren Zusammenhang auf der Fläche man in einem Streckenkomplex wiedergibt. Der Streckenkomplex spiegelt die Verheftung der Halbblätter über gewissen Teilbogen von L wider: zu einem bestimmten Streckenkomplex gehört bei gegebener Zerschneidungskurve und darauf gegebenen Grundpunkten stets eine bis auf zyklische Vertauschung der Grundpunkte wohlbestimmte RIEMANNsche Fläche. $\mathfrak{M}$ sei die Klasse aller einfach zusammenhängenden Flächen, deren Windungspunkte sich in endlich viele Grundpunkte projizieren.

Der Streckenkomplex erwies sich in vielen Problemen der Wertverteilungslehre als ein sehr kräftiges Hilfsmittel, denn er gestattet, jede Fläche $M \in \mathfrak{M}$ durch einen Streckenkomplex in übersichtlicher Weise darzustellen. Allerdings versagt dieses Hilfsmittel zunächst bei RIEMANNschen Flächen, deren Verzweigungsstellen nicht mehr über endlich vielen Grundpunkten liegen, denn dieser Streckenkomplexbegriff S ist auf die Klasse $\mathfrak{M}$ zugeschnitten; andererseits sind die nicht über endlich vielen Grundpunkten verzweigten Flächen natürlich von besonderem Interesse für die Theorie der Wertverteilung: eine Fläche $M \in \mathfrak{M}$ hat als Randstellen beispielsweise nur logarithmische Windungspunkte. Nun gibt der Streckenkomplex nur eine topologische Darstellung des Aufbaus einer RIEMANNschen Fläche aus *Halbblättern*; man wird also fragen, ob es denn nicht möglich sei, auch solche Flächen durch einen Streckenkomplex darzustellen, die zwar die Eigenschaft, nur über endlich vielen Punkten verzweigt zu sein, nicht haben, sich aber immer noch aus Halbblättern aufbauen, also Flächen sind, deren Verzweigungspunkte und Randstellen über einer geschlossenen doppelpunktfreien JORDAN-Kurve L

*) Die Arbeit wurde von der Deutschen Forschungsgemeinschaft gefördert; ich sage dafür meinen Dank.

1) NEVANLINNA [4], [5] Kap. XI § 2; ELFVING [1]. Man verwendet die Konstruktion hauptsächlich bei der Unterklasse der einfach zusammenhängenden Flächen aus der Klasse der schlichtartigen Flächen.

der Ebene liegen. Es zeigt sich, daß mit einer geeigneten Definition des Streckenkomplexes, die von der bisher gegebenen in einigen Punkten abweicht, die Frage positiv zu beantworten ist.

§ 2. Auch in der Vergangenheit wurden schon Abänderungen und Verallgemeinerungen des Streckenkomplexbegriffes vorgenommen, um den Bereich der Flächen zu erweitern, die sich in einem Streckenkomplex darstellen lassen: Man ordnete jedem Grundpunkt a_ν $(\nu = 1, 2, \ldots, n)$ einer nur über diesen Punkten verzweigten Fläche M ein einfach zusammenhängendes Gebiet G_ν der Art zu, daß $a_\nu \in G_\nu$ und daß zwei verschiedene Gebiete G_ν, G_μ positiven Abstand voneinander haben; schiebt man nun einen Windungspunkt über a_ν in einen beliebigen Punkt $a'_\nu \in G_\nu$ — wobei verschiedene Windungspunkte der Sorte a_ν auch über verschiedene Punkte geschoben werden können — so entsteht aus M eine Fläche M'; dieser wurde der Streckenkomplex der Fläche M zugeordnet. Eine solche Verwendung des Streckenkomplexes findet sich mehrfach in der letzten Zeit.

Eine weitere naheliegende Erweiterung bestände darin, zuzulassen, daß ein Gebiet G_ν mit einem anderen G_μ einen Berührungspunkt gemeinsam habe. Eine Begriffsbildung dieser Art wurde schon von ULLRICH in Vorlesungen und Vorträgen verwendet, sie findet sich in der Literatur später bei seinem Schüler HANCK[2]) und bei TEICHMÜLLER[2]), jedoch in einer Weise, die das Hauptaugenmerk auf bestimmte Anwendungsmöglichkeiten richtet. Dies bedingte, daß die eigentliche Definition eines erweiterten Streckenkomplexbegriffes unterblieb, denn der Blick auf die Anwendung brachte es mit sich, daß der erweiterte Begriff nur so weit beleuchtet wurde, als für die Anwendung erforderlich war.

§ 3. Es erscheint nunmehr natürlich, einen erweiterten Streckenkomplexbegriff axiomatisch einzuführen. Eine Anregung zu dieser Arbeit verdanke ich Gesprächen mit Miss M. L. CARTWRIGHT über diesen Gegenstand, in denen mir das Fehlen einer allgemeinen Definition für den erweiterten Begriff des Streckenkomplexes deutlich wurde. Es zeigt sich, daß eine geeignete Definition eine weit umfassendere Flächenklasse der streckenkomplexmäßigen Darstellung zugänglich macht, als man nach den bisherigen Erweiterungen vielleicht erwartet hätte: Nicht nur ist es nicht erforderlich, daß überhaupt nur über endlich vielen Punkten Verzweigungen vorkommen, sondern es ist nicht einmal notwendig, daß am Rand der einzelnen Halbblätter nur endliche viele Verzweigungspunkte oder Randstellen liegen.

Ich beschränke mich auf diejenigen einfach zusammenhängenden Flächen, die von Funktionen erzeugt werden, welche als Träger eines der drei schlichten Normalgebiete (Einheitskreis, punktierte Ebene, Vollebene) haben, wobei im Fall des Einheitskreises als Träger die Funktion auf keinen Teilbogen des Einheitskreises stetig fortsetzbar sei (so daß also kein JORDAN-Bogen ein Randstück der Fläche bildet).

Zur sprachlichen Vereinfachung will ich die Windungspunkte und Randstellen einer RIEMANNschen Fläche unter den umfassenden Begriff *Singu-*

[2]) HANCK [2], TEICHMÜLLER [7].

laritäten der Fläche unterordnen. Endlich-vielblättrige Windungspunkte sind algebraische Singularitäten; alle übrigen Singularitäten sind transzendent, sie sind insbesondere also keine (inneren) Punkte der Fläche.

§ 4. Ich betrachte am Anfang die topologische Abbildung einfach zusammenhängender RIEMANNscher Flächen, deren Singularitäten nur über einer geschlossenen doppelpunktfreien JORDAN-Kurve L liegen, in die schlichte Ebene (§ 5). Von dort ergeben sich die charakteristischen Eigenschaften des Streckenkomplexes S^*, welche der axiomatischen Einführung des Begriffes zugrunde liegen (§§ 6—9); ebenso wird man auf die Vorschrift geführt, die über die Zuordnung von Streckenkomplex $S^*(W)$ und RIEMANNscher Fläche W zu geben sind (§ 10). Danach wende ich mich der Frage zu, welche Gestalt ein Streckenkomplex S^* haben muß, damit er eine einfach zusammenhängende Fläche darstellt (§ 11); hier wird sich eine notwendige und hinreichende Bedingung ergeben (Satz 1). Eine topologische Abbildung einer im Streckenkomplex dargestellten RIEMANNschen Fläche (§ 12) erlaubt, ein genaues Entsprechen von Randstellen der Fläche und gewissen Strukturen des Streckenkomplexes zu erkennen (§§ 13—19, Sätze 2, 3). Zum Schluß werde ich ein paar typische Beispiele von Flächen geben, die sich nun im Streckenkomplex darstellen lassen (§ 20), und den erweiterten Streckenkomplexbegriff dem bisherigen gegenüberstellen (§ 21).

Konstruktion des Streckenkomplexes und seine Eigenschaften

§ 5. Gegeben sei eine einfach zusammenhängende RIEMANNsche Fläche W als Überlagerung der w-Ebene; ihre sämtlichen Singularitäten mögen über einer geschlossenen doppelpunktfreien JORDAN-Kurve L liegen, so daß also in jedem Punkt P über $p \notin L$ entweder $w-p$ oder w^{-1} Ortsuniformisierende ist. Es werde ferner angenommen, daß in keinem Blatt der Fläche ein Teilbogen von L die Fläche berande. Ausstanzen über L zerlegt W in Halbblätter, die auf W über gewisse Teilbogen von L hinweg zusammenhängen. Sie werden als *Innengebiete* bzw. *Außengebiete* bezeichnet, je nachdem sie zur Linken oder zur Rechten liegen, wenn L in bestimmtem Sinn orientiert wird. Wird W durch konforme Abbildung $z(w)$ in das schlichte Normalgebiet — die Vollebene, die punktierte Ebene $z \neq \infty$ oder den Einheitskreis $|z|<1$ — übergeführt, so erhält man in den Bildern der Halbblätter eine aus Kurvenpolygonen, den *Elementarpolygonen*, gebildete Figur F, die die Verheftung der Halbblätter widerspiegelt. Jede mit F topologisch äquivalente Figur in der Ebene wird als *Netz* N der Fläche W bezeichnet; N entsteht aus W durch eine gewisse topologische Abbildung $h(w)$, umgekehrt W aus N durch die eindeutige Funktion $h^{-1}(z)$. Die Gesamtheit der Randpunkte der Figur F bzw. des topologischen Bildes N heißt der *Rand* des Netzes; die Bilder der über L liegenden Kurven heißen *Netzkurven* oder *Netzlinien*, sie sind Berandungslinien der Elementarpolygone. Ist die Fläche W nicht-kompakt — ist also das schlichte Normalgebiet die punktierte Ebene oder der Einheitskreis — und wird F als Netz gewählt, so ist natürlich der Punkt $z=\infty$ bzw.

die Kreislinie $|z|=1$ der Rand des Netzes. Der Schnittpunkt P von n Netzlinien ($2 \leq n < \infty$) ist vermittels $h(w)$ Bild eines algebraischen Windungspunktes, in dem n Blätter zusammenhängen; in P treffen $2n$ Elementarpolygone zusammen, solche Punkte nennen wir *Eckpunkte* des Netzes. Ist $N=F$, so ist P ein Kreuzungspunkt der zu $z(w)$ inversen meromorphen Funktion $w(z)$, die die Fläche W erzeugt. Ein keinen Eckpunkt im Innern enthaltendes Stück s einer Netzkurve zwischen zwei Eckpunkten, einem Eckpunkt und dem Rand, oder zwischen Randpunkten des Netzes nennen wir eine *Seite* des Netzes. Ist $N=F$, so ist s Bild eines offenen Regularitätsbogens von $z(w)$. Zwei Elementarpolygone heißen *benachbart*, wenn sie eine Netzseite gemeinsam haben.

Eine Folge von Punkten lauter verschiedener Netzseiten häuft sich nur gegen den Rand des Netzes; wir drücken dies kurz so aus: *Die Seiten des Netzes häufen sich nur gegen seinen Rand.* Zum Beweis dieser topologischen Behauptung darf $N=F$ angenommen werden. Es sei z_0 ein Häufungspunkt von Punkten aus lauter verschiedenen Netzseiten. Dann gibt es entweder eine Folge z_ν von Eckpunkten des Netzes [also von Kreuzungspunkten der Funktion $w(z)$] der Art, daß $\lim z_\nu = z_0$ — in diesem Fall ist z_0 wegen $w'(z_\nu)=0$ kein innerer Punkt des schlichten Normalgebietes F — oder es gibt ganze abgeschlossene Stücke l_ν auf Netzseiten s_ν, so daß ebenfalls $z_\nu \in l_\nu$ und $\lim z_\nu = z_0$, während Eckpunkte am Ende der Seiten s_ν den Punkt z_0 nicht als Häufungspunkt haben — aber auch diese Annahme schließt aus, daß z_0 ein innerer Punkt von F sei, denn eine (hinreichend kleine) Umgebung eines solchen enthält höchstens von endlich vielen Netzseiten Stücke.

§ 6. Die Verheftung der Halbblätter überträgt sich auf das Aneinanderstoßen der Elementarpolygone im Netz und kann in folgender Weise durch einen *Streckenkomplex* dargestellt werden: In jedem Elementargebiet wird ein Punkt, *Knoten*, markiert, mit einem Ring in den Bildern von Innengebieten, mit einem Kreuz in den Bildern von Außengebieten. Wie die Netzseiten häufen sich die Knoten nur gegen den Rand des Netzes. Jeder Ring wird mit den Kreuzen in den benachbarten Elementarpolygonen, jedes Kreuz mit den benachbarten Ringen durch JORDAN-Bögen, *Glieder*, in der Weise verbunden, daß

(1a) *jede Netzseite genau einmal (in einem inneren Punkt) geschnitten wird,*

(1b) *außer in den Ringen und Kreuzen keine Überschneidungen der Glieder auftreten.*

In einem Knoten laufen k Glieder zusammen, wenn k die Anzahl der (Netz-)Seiten am Rand des Elementarpolygons ist, in dem der Knoten liegt; es ist $k \geq 2$, weil kein Teilbogen von L die Fläche W berandet, k kann jedoch unendlich sein. Die nach (1a), (1b) erzeugte Figur $S^*(W)$ heißt *Streckenkomplex* der Fläche W[3]).

[3]) Im Gegensatz zum Streckenkomplex S kommen im Streckenkomplex S^* keine Mehrfachbindungen vor: zwei Knoten sind nicht durch benachbarte Glieder verbunden.

Mitunter ist es zweckmäßig, bei der Konstruktion des Streckenkomplexes die Bedingung (1a) zu verschärfen: Jedem im Sinn der Orientierung auf L von a nach b durchlaufenen offenen Teilbogen $I(a, b)$ werde ein Punkt $l(a, b) \in I(a, b)$ so zugeordnet, daß $l(a, b)$ in beiden Argumenten stetig ist und daß mit $c \in I(a, b)$ auch $l(a, c) \in I\{a, l(a, b)\}$ und $l(c, b) \in I\{l(a, b), b\}$; dies bedeutet eine gewisse Monotonieeigenschaft der Funktion $l(a, b)$ in beiden Argumenten. Diese Zuordnung ist möglich: Ist L' die Kreislinie $|w| = 1$, orientiert in Richtung wachsenden Arguments, so wähle man mit $0 < \psi - \varphi < 2\pi$ einfach $l(e^{i\varphi}, e^{i\psi}) = e^{\frac{1}{2}i(\varphi+\psi)}$ und übertrage die Zuordnung durch topologische Abbildung auf die gegebene Kurve L. Eine Verschärfung von (1a) ist die zu (1a) zusätzliche Forderung

(1a') *Eine Netzseite, die vermittels $h(w)$ Bild des Teilbogens $I(a, b) \subset L$ ist, werde vom Glied des Streckenkomplexes im Punkt $h\{l(a, b)\}$ geschnitten.*

§ 7. Aus (1a), (1b) folgt, daß ein innerer Punkt eines Gliedes nie Häufungspunkt von Punkten anderer Glieder ist. Es folgt insbesondere, daß Häufungsmengen von Gliedern JORDAN-Bögen bilden, die ohne eine Netzseite zu treffen von Knoten aus zum Rand des Netzes führen und, eventuell von Knoten abgesehen, im Innern des Netzes keine gemeinsamen Punkte haben. [Man bilde ein Elementarpolygon mit unendlich vielen Seiten topologisch ab auf einen Kreis, in dem die Bilder der Glieder auf Radien liegen.] Wir nennen diese JORDAN-Bögen *Strahlen* des Streckenkomplexes S^*. Jeder Strahl liegt also in einem wohlbestimmten Elementarpolygon, und die Menge der Strahlen ist nirgends dicht.

Diejenigen keinen Knoten im Inneren enthaltenden Teilgebiete der Ebene, in die die Glieder und die Strahlen des Streckenkomplexes S^* das Netz zerlegen, heißen *Elementargebiete*[4]; die Glieder und Strahlen am Rand eines Elementargebietes heißen seine *Seiten*, die am Rand liegenden Knoten seine *Eckknoten*. Weil jeder Strahl einen Knoten mit dem Rand des Netzes verbindet, können höchstens zwei Strahlen ein Elementargebiet beranden; ein Elementargebiet, das genau einen Strahl als Seite besitzt, hat unendlich viele Seiten, weil die Gliederfolge seines Randes gegen den Netzrand streben muß; ein Elementargebiet mit zwei verschiedenen Strahlen als Seiten hat dagegen nur endlich viele Seiten, weil die Knoten, die zu den beiden Strahlen gehören, über eine endliche Zahl von Gliedern miteinander verbunden sind. Ein Elementargebiet mit zwei Strahlen am Rand nennen wir *Strahl-Elementargebiet*. Die Elementargebiete, deren Rand eine endliche — und somit gerade — Anzahl $2n$ von *Gliedern* bildet, entsprechen den Eckpunkten des Netzes, also den n-blättrigen algebraischen Windungspunkten der Fläche W; danach heißen sie *algebraische Elementargebiete*[5]. [Es ist stets $n \geqq 2$, weil für $n = 1$ ja auf W keine Singularität vorliegt, also kein Eckpunkt des Netzes entstehen kann, vgl. Fußnote [3].] Die Koordinate a des einem algebraischen Elementargebiet

[4] Elementargebiete können Randpunkte des Netzes am Rand enthalten.

[5] Ein endlich-vielseitiges Elementargebiet im Streckenkomplex S^* ist nicht notwendig ein algebraisches Elementargebiet.

entsprechenden algebraischen Windungspunktes auf W heißt *Koordinate* des Elementargebietes; dem Elementargebiet ist die komplexe Zahl a *zugeordnet*. Ein Elementargebiet mit unendlich vielen Seiten werde in Anlehnung an die Terminologie, wie sie im Zusammenhang des Streckenkomplexes S üblich ist, *logarithmisches Elementargebiet* genannt[6]); es hat mit dem Rand des Netzes (wenigstens) einen Randpunkt gemeinsam, weil die Netzseiten sich nicht im Innern des Netzes häufen; ebenso hat trivialerweise ein Strahl-Elementargebiet mit dem Rand des Netzes einen Randpunkt gemeinsam. Ein logarithmisches Elementargebiet bedeckt ein Gebiet der z-Ebene, in dem alle Netzlinien, ein Strahl-Elementargebiet bedeckt ein Gebiet, in dem überhaupt alle Wege, die zum Rand des Netzes laufen, für $h^{-1}(z)$ Zielwege mit demselben Zielwert sind. Der Zielwert a ist in allen Halbblättern, deren Knoten Eckknoten des (logarithmischen oder Strahl-) Elementargebietes sind, eine Singularität der Fläche W, und natürlich eine transzendente; das Elementargebiet ist dieser transzendenten Singularität zugeordnet, wie ein algebraisches Elementargebiet einer algebraischen Singularität der Fläche zugeordnet ist. Die Zahl a heißt *Koordinate* des (logarithmischen oder Strahl-) Elementargebietes; sie ist ihm *zugeordnet*. Ein Strahl-Elementargebiet hat zwei Strahlen des Streckenkomplexes als Seiten, jeder ist einer Häufungsstelle von Singularitäten am Rand eines Halbblattes zugeordnet; eine kleine Überlegung zeigt, daß die Koordinaten beider Häufungsstellen gleich sind[7]).

Jedem Glied des Streckenkomplexes, das einen Ring mit einem Kreuz verbindet, entspricht ein bestimmter Teilbogen von L, über den die den beiden Knoten zugehörigen Halbblätter verheftet sind; weil von jedem Knoten wenigstens zwei Glieder ausgehen, trennt jedes Glied zwei Elementargebiete, und zwar die beiden, die den beiden Singularitäten am Ende des Teilbogens zugeordnet sind. Zwei Glieder, die beim Umlauf um einen Knoten unmittelbar aufeinander folgen, heißen *benachbart*; ebenso heißen zwei Elementargebiete, die durch ein Glied getrennt werden, benachbart. Die Glieder des Streckenkomplexes S^* können sich, wie wir sahen — anders als die Seiten des Netzes — im Innern des Netzes auf Strahlen häufen; dies tritt genau dann ein, wenn die RIEMANNsche Fläche W Halbblätter enthält, auf deren Rand unendlich viele Singularitäten liegen. Den Häufungspunkten der Singularitäten entsprechen Häufungsmengen von Gliedern und somit von Elementargebieten.

§ 8. Wir schließen im folgenden den im wesentlichen bekannten Fall aus, die RIEMANNsche Fläche W enthalte nur endlich viele Blätter; sie würde als schlichtartige Fläche nur über endlich vielen Punkten verzweigt sein, also schon durch den Streckenkomplex S darstellbar sein. *Wir nehmen ferner an,*

[6]) Im Gegensatz zu den Verhältnissen beim Streckenkomplex S braucht einem logarithmischen Elementargebiet im Streckenkomplex S^* als Singularität kein logarithmischer Windungspunkt zuzugehören; auch kann ein Strahl Seite eines logarithmischen Elementargebietes sein.

[7]) Im § 19 werden wir sehen, daß beide Häufungsstellen dieselbe Randstelle definieren.

der unendlich ferne Punkt U der z-Ebene sei der einzige Randpunkt des Netzes; dies läßt sich durch topologische Abbildung stets erreichen.

§ 9. Der Streckenkomplex S^* hat nun offenbar folgende Eigenschaften:

1°. *Die Anzahl der Knoten ist abzählbar unendlich und der Punkt U ist der einzige Häufungspunkt der Knoten.*

2°. *Jedes Glied verbindet einen Ring mit einem Kreuz.*

3°. *Verschiedene Glieder haben außer den Knoten keine gemeinsamen Punkte.*

4°. *Der Streckenkomplex ist zusammenhängend: zwei Knoten sind stets über eine Gliederfolge verbunden.*

5°. *Zwei Knoten sind nicht durch benachbarte Glieder verbunden*[8]).

6°. *Von jedem Knoten gehen mindestens zwei Glieder aus.*

7°. *Kein von einem Knoten verschiedener Punkt eines Gliedes ist Häufungspunkt von Punkten anderer Glieder.* Es folgte

a) *jedes Glied trennt zwei Elementargebiete;*

b) *Häufungsmengen von Gliedern liegen nur auf punktfremden* JORDAN-*Kurven, den Strahlen, die von einem Knoten aus zum Punkt U führen.*

8°. *Jedem Elementargebiet E_ν ist eine komplexe Zahl a_ν zugeordnet.*

9°. *Zwei benachbarten Elementargebieten sind verschiedene Zahlen zugeordnet.*

10°. *Wird ein Ring R genau einmal in positivem Sinn umlaufen, so folgen die den Elementargebieten mit dem Eckknoten R zugeordneten Zahlen in solcher Weise aufeinander, daß die Kurve L genau einmal im Sinn der Orientierung durchlaufen wird; dasselbe gilt beim Umlauf um ein Kreuz in negativem Sinn.*

11°. *Jedem Strahl σ_γ ist eine komplexe Zahl a_γ zugeordnet, nämlich der Grenzwert der Zahlen, die den sich dort häufenden Elementargebieten zugeordnet sind.* [Denn eine Häufungsmenge der Glieder ist auch eine Häufungsmenge von Elementargebieten, und der Grenzwert der Zahlen, die einer gegen σ_γ konvergierenden Folge von Elementargebieten zugeordnet sind, existiert und ist unabhängig von der Folge (vgl. die Ordnungseigenschaft 10°).]

12°. *Der zur Zahl a_ν* bzw. *a_γ gehörige Punkt der Zahlenebene liegt auf einer geschlossenen doppelpunktfreien* JORDAN-*Kurve L.*

13°. *Ein innerer Punkt G^μ eines Gliedes g_ν^μ, das ein Elementargebiet E_ν berandet, kann mit einem inneren Punkt T_ν von E_ν falls E_ν algebraisches, mit dem Randpunkt ∞ von E_ν falls E_ν nicht algebraisches Elementargebiet ist*[9]), *so durch einen — bis eventuell auf den Endpunkt — in E_ν verlaufenden* JORDAN-*Bogen s_ν^μ verbunden werden, daß die Gesamtheit der Bögen s_ν^μ bis auf die Punkte G^μ, T_ν punktfremd sind und sich nicht im Endlichen häufen*[10]).

Die einem Elementargebiet bzw. einem Strahl zugeordnete komplexe Zahl heißt *Koordinate* des Elementargebietes bzw. des Strahles. Der einfacheren

[8]) Vgl. Fußnote [3]).

[9]) Nach 1°, 7°b ist im letzten Fall U ein Randpunkt des Elementargebietes, nach 4° hat es nur einen Randpunkt bei U.

[10]) Wir sagen: Ein Punkt P ist *Häufungspunkt* einer unendlichen Menge von JORDAN-Bögen (Gebieten), wenn es eine Folge von Punkten aus lauter verschiedenen JORDAN-Bögen (Gebieten) gibt, die den Punkt P als Häufungspunkt hat.

Sprechweise wegen bezeichnen wir im folgenden die Elementargebiete und ihre Häufungsmengen einfach auch mit den zugehörigen Koordinaten.

Die Eigenschaften 1° bis 13° sind charakteristisch für den Streckenkomplex S^* einer einfach zusammenhängenden RIEMANNschen Fläche[11]).

Zuordnung der RIEMANNschen Fläche zum Streckenkomplex

§ 10. Es sei nämlich eine orientierte geschlossene doppelpunktfreie JORDAN-Kurve L gegeben, die die Ebene in zwei Gebiete, die Halbblätter, zerlegt, sowie ein Streckenkomplex S^*, der die Bedingungen 1° bis 13° erfüllt. Wir ordnen ihm eine RIEMANNsche Fläche W in der folgenden Weise zu:

Jedem Ring wird das bei positiver Durchlaufung von L links liegende Gebiet, das *Innengebiet*, jedem Kreuz das rechts liegende, das *Außengebiet*, zugeordnet; sind zwei Knoten durch ein Glied verbunden, das die Elementargebiete a_i, a_k trennt (cf. 7°a, 8°), so werden die zugehörigen Halbblätter — es sind nach 2° ein Innen- und ein Außengebiet — über demjenigen Teilbogen s_{ik} [$= I(a_i, a_k)$ bzw. $I(a_k, a_i)$ nach § 6] von L zwischen a_i, a_k (cf. 12°) verheftet, dessen Durchlaufung dem Überschreiten des Gliedes (cf. 10°) entspricht; es ist $a_i \neq a_k$ (cf. 9°). Wir nennen den Bogen s_{ik} *verheftenden* Teilbogen.

Auf diese Weise ist dem Streckenkomplex S^* eine zusammenhängende (cf. 4°) offene zweidimensionale Mannigfaltigkeit W zugeordnet, die eine RIEMANNsche Fläche ist: denn in einem Punkt P über der endlichen Koordinate a (bzw. der Koordinate ∞) in einem Innen- oder Außengebiet oder auf einem verheftenden Teilbogen ist $w - a$ (bzw. w^{-1}) Ortsuniformisierende; ein Endpunkt P eines verheftenden Teilbogens werde genau dann zur Fläche hinzugenommen, wenn er zu einem algebraischen Elementargebiet gehört. Nach 2° wird ein solches Elementargebiet von einer geraden Zahl $2n$ von Gliedern berandet [$n \geqq 2$ (cf. 5°)], ein n-maliger Umlauf um P führt also ins ursprüngliche Halbblatt zurück. Hat P die endliche Koordinate a (bzw. die Koordinate ∞), so ist $(w-a)^{1/n}$ (bzw. $w^{-1/n}$) Ortsuniformisierende in P. P ist also ein algebraischer Windungspunkt, in dem $n \geqq 2$ Blätter zusammenhängen. Die Nachbarrelation zwischen Ortsuniformisierenden, deren Definitionsgebiete einen nicht leeren Durchschnitt haben, ist offenbar direkt konform.

Logarithmische Elementargebiete, Strahl-Elementargebiete und Strahlen fassen wir unter den Begriff *Randelement* des Streckenkomplexes S^*[12]); einem Randelement ist zwar formal eine komplexe Zahl a zugeordnet (nach 8° oder 11°), ihm entspricht aber kein (innerer) Punkt der Fläche W, sondern eine *Randstelle* von W über a[13]).

[11]) Die Eigenschaft 1° ist natürlich einfach eine Folge der Verabredung von § 8, die Eigenschaft 5° ist mehr formaler Natur, vgl. dazu § 21, Fußnote [23]). Die Anzahl der Elementargebiete ist abzählbar, denn jeder Knoten ist nach 7° Eckknoten von höchstens abzählbar vielen Elementargebieten. Die Anzahl der Strahlen kann dagegen überabzählbar sein.

[12]) Unter diesen Begriff fallen noch andere Elemente des Streckenkomplexes; vgl. dazu § 16.

[13]) Es ist dabei möglich, daß verschiedenen Randelementen von S^* dieselbe Randstelle von W entspricht. Genaueres in §§ 17—19.

§ 11. Es gilt nun der

Satz 1. *Erfüllt der Streckenkomplex die Bedingungen* 1° *bis* 13°, *so ist die zugeordnete* RIEMANNsche *Fläche W (schlichtartig) und einfach zusammenhängend. Umgekehrt hat jede unendlich-vielblättrige einfach zusammenhängende* RIEMANNsche *Fläche, die nur über einer geschlossen doppelpunktfreien* JORDAN-*Kurve L Singularitäten trägt und von keinem Teilbogen von L berandet wird, einen Streckenkomplex S*, der die Bedingungen* 1° *bis* 13° *erfüllt.*

Der zweite Teil des Satzes ist das bereits in § 9 gewonnenes Ergebnis. Zum Beweis des ersten Teiles weisen wir zunächst nach, daß jeder Rückkehrschnitt die Fläche W zerlegt, daß W also schlichtartig ist. Der einfache Zusammenhang von W folgt dann daraus, daß einer der beiden Teile, in die ein Rückkehrschnitt die Fläche zerlegt, kompakt ist[14]). Auf eine topologische Deformation des Rückkehrschnittes kommt es hierbei nicht an[15]).

Wir bilden jeden Knoten K_n von S^* ab auf einen bestimmten *festen* inneren Punkt P_i bzw. P_a des zugeordneten Innen- bzw. Außengebietes, und jedes Glied zwischen zwei Knoten topologisch ab auf eine Kurve, die nach Wahl einer Funktion $l(a, b)$ die beiden in den zugehörigen Halbblättern ausgezeichneten Punkte P_i, P_a über den inneren Punkt $l(a, b)$ des verheftenden Teilbogens $I(a, b)$ miteinander verbindet [cf. § 6] und mit diesem nur den Punkt $l(a, b)$ gemein hat. Die Abbildung der Glieder geschehe in der Weise, daß die Projektionen zweier solcher Bildkurven in die w-Ebene genau dann zusammenfallen, wenn sie den verheftenden Teilbogen über demselben Punkt von L treffen, während sie sonst nur zwei Punkte, nämlich P_i und P_a, gemeinsam haben. So erhalten wir im Streckenkomplex S^* das topologische Bild eines gewissen Kurvenzuges auf W.

Sei $\mathfrak{r}$ ein Rückkehrschnitt auf W. Er ist kompakt und kann als Polygonzug angenommen werden. Wir zeigen zunächst, daß $\mathfrak{r}$ nur durch endlich viele Halbblätter läuft. Die Annahme nämlich, $\mathfrak{r}$ liefe durch unendlich viele, würde erlauben, in jedem dieser Halbblätter H_ν genau einen Punkt w_ν von $\mathfrak{r}$ zu markieren. Wegen der Kompaktheit von $\mathfrak{r}$ muß die Punktfolge w_ν einen Häufungspunkt P der Koordinate p auf W haben. In der Umgebung von P würden unendlich viele Blätter zusammenhängen, andererseits ist (mit ganzzahligem n) entweder $(w-p)^{1/n}$ oder $w^{-1/n}$ Ortsuniformisierende in P, eine hinreichend kleine Umgebung von P enthält also Punkte aus höchstens n Blättern. Aus dem Widerspruch folgt: die Annahme oben war irrig.

Wir zeigen nun, daß $\mathfrak{r}$ die Fläche W zerlegt. Es kann abgesehen werden von dem Fall, daß $\mathfrak{r}$ durch einen Windungspunkt hindurchläuft. [Dann kann nämlich $\mathfrak{r}$ in der Umgebung des Windungspunktes in eine Kurve topologisch deformiert werden, die den Windungspunkt in geeigneter Weise umschlingt[15]).] Wir dürfen aber sogar annehmen, daß $\mathfrak{r}$ eine *einfache Folge* H_ν von Halbblättern, d.h. jedes nur einmal, durchläuft, also daß der Durchschnitt von $\mathfrak{r}$ mit H_ν zusammenhängend ist. Um dies einzusehen, stellen wir $\mathfrak{r}$ in Parameterform

[14]) Der Beweis schließt sich Gedankengängen von ELFVING [1], S. 11f., an.

[15]) Zur topologischen Deformation von Wegen s. z.B. NEVANLINNA [6], S. 59.

als $\mathfrak{r}(t)$, $0 \leq t \leq 1$, dar; die Punkte $\mathfrak{r}(0)$ und $\mathfrak{r}(1)$ und nur diese fallen zusammen. Nehmen wir an, es gebe im Intervall $(0, 1)$ zwei Zahlen t', t'' $(t' < t'')$, so daß $\mathfrak{r}(t') \in H_\nu$, $\mathfrak{r}(t'') \in H_\nu$, $\mathfrak{r}(0) \notin H_\nu$ aber $\mathfrak{r}(t^*) \notin H_\nu$, wo $t' < t^* < t''$, so verbinden wir $\mathfrak{r}(t')$ mit $\mathfrak{r}(t'')$ durch eine in H_ν verlaufende doppelpunktfreie *Querkurve* $\bar{\mathfrak{r}}(t)$, $t' \leq t \leq t''$, die mit $\mathfrak{r}$ nur die Punkte $\mathfrak{r}(t') = \bar{\mathfrak{r}}(t')$ und $\mathfrak{r}(t'') = \bar{\mathfrak{r}}(t'')$ gemein hat. $\mathfrak{r}(t)$ und $\bar{\mathfrak{r}}(t)$ bestimmen zusammen zwei verschiedene Rückkehrschnitte, die nur dann beide die Fläche zerlegen, wenn $\mathfrak{r}$ schon diese Eigenschaft hatte.

Weiter dürfen wir annehmen, daß $\mathfrak{r}$ nicht über ein und denselben verheftenden Teilbogen b_ν in das Halbblatt H_ν eintritt und es verläßt, denn in diesem Fall bildet ein Teilbogen von b_ν eine Querkurve, die $\mathfrak{r}$ in zwei Teile zerlegt, von denen einer, nämlich der in und am Rand von H_ν verlaufende, die Fläche gewiß zerlegt. Als geschlossene Kurve, die eine einfache Folge von Halbblättern durchläuft und einen verheftenden Teilbogen nicht zweimal trifft, ist $\mathfrak{r}$ topologisch äquivalent entweder mit einem in einem einzigen Halbblatt liegenden Kreis — dann zerlegt $\mathfrak{r}$ gewiß die Fläche — oder mit dem Bild einer endlich viele Knoten genau einmal treffenden geschlossenen doppelpunktfreien (cf. 3°) Kurve $\mathfrak{s}$ aus Gliedern von S^*, in der jedes überhaupt durchlaufene Glied genau einmal in bestimmter Richtung durchlaufen wird. Jede doppelpunktfreie Kurve $\bar{\mathfrak{s}}$ aus Gliedern von S^*, die zwischen zwei verschiedenen Knoten von $\mathfrak{s}$ verläuft, und bis auf diese ihre Endpunkte innerhalb des von $\mathfrak{s}$ eingeschlossenen Gebietes G, hat als Bild auf W eine Querkurve $\bar{\mathfrak{r}}$ von $\mathfrak{r}$; $\mathfrak{s}$ umschließt als Urbild einer Kurve, die nur endlich viele Halbblätter trifft, endlich viele algebraische Elementargebiete (cf. 1°), es gibt daher nur endlich viele Gliedzugkurven $\bar{\mathfrak{s}}$ der betrachteten Art. Die Zerlegung von $\mathfrak{s}$ durch, bis eventuell auf die Endpunkte, ganz innerhalb G verlaufende Gliedzugkurven $\bar{\mathfrak{s}}$ führt schließlich auf die Ränder von endlich vielen algebraischen Elementargebieten. Nun ist eine Gliedzugkurve $\mathfrak{s}^*$, die ein $2n$-eckiges algebraisches Elementargebiet umschließt, offenbar auf eine Kurve $\mathfrak{r}^*$ auf W abgebildet, die einen n-blättrigen Windungspunkt einmal umschlingt; eine solche Kurve zerlegt die Fläche. Es zerlegt also auch $\mathfrak{r}$ die Fläche W; damit ist die Schlichtartigkeit von W bewiesen.

Es bleibt noch zu zeigen, daß von den zwei Teilen W_1, W_2, in die $\mathfrak{r}$ die Fläche W zerlegt, der eine kompakt ist. Dies ist offenbar gewiß dann der Fall, wenn $\mathfrak{r}$ eine Kurve $\mathfrak{r}^*$ ist, die einen algebraischen Windungspunkt einmal umschlingt, oder eine Kurve, die in *einem* Halbblatt verläuft. Im allgemeinen Fall dürfen wir wieder annehmen, $\mathfrak{r}$ durchlaufe eine einfache Folge von Halbblättern und trete ein in ein Halbblatt und verlasse es nicht über denselben verheftenden Teilbogen, denn die beiden Rückkehrschnitte, die durch $\mathfrak{r}$ und eine Querkurve $\bar{\mathfrak{r}}$ bestimmt werden, umschließen nur dann beide kompakte Teilstücke von W, wenn $\mathfrak{r}$ schon diese Eigenschaft hat. Wir betrachten die zu $\mathfrak{r}$ topologisch äquivalente Gliedzugkurve $\mathfrak{s}$ und die Gesamtheit der oben eingeführten Gliedzugkurven $\bar{\mathfrak{s}}$, die das Gebiet G in endlich viele algebraische Elementargebiete zerlegen. Seien S_1, S_2 zwei Punkte auf diesen Gliedzugkurven, die nicht auf $\mathfrak{s}$ liegen. Die zugehörigen Bildpunkte R_1, R_2 auf W werden durch $\mathfrak{r}$ nicht getrennt: S_1, S_2 sind ja durch Teile von Gliedzugkurven

verbunden, die $\mathfrak{r}$ nicht überschreiten. Demnach liegen die Bilder aller nicht auf $\mathfrak{s}$ liegenden Punkte der Kurven $\bar{\mathfrak{s}}$ in ein und demselben der beiden Teilgebiete, in die $\mathfrak{r}$ die Fläche W zerlegt, etwa in W_1. W_1 wird selbst durch die Bilder $\bar{\mathfrak{r}}$ der Kurven $\bar{\mathfrak{s}}$ in endlich viele Teilgebiete zerlegt, die nach Hinzunahme der Punkte auf den Kurven $\bar{\mathfrak{r}}$ sämtlich kompakte Teilflächenstücke von W sind, denn sie sind abgeschlossene Umgebungen algebraischer Windungspunkte. Es ist also auch W_1 kompakt; damit ist der einfache Zusammenhang von W bewiesen.

§ 12. Vom Streckenkomplex S^* gelangen wir zum Netz N von W in der folgenden Weise. In jedem Elementargebiet E_ν der Koordinate a_ν, das am Rand m Glieder g_ν^μ $(1 \leq m \leq \infty)$ besitzt, wird ein Punkt T_ν markiert — seine Lage wird später genauer festgelegt —; von ihm aus werden in E_ν m JORDAN-Kurven s_ν^μ $(\mu = 1, 2, \ldots, m)$ gezogen, und zwar werde jedes Randglied g_ν^μ von E_ν durch genau eine solche Kurve s_ν^μ mit T_ν verbunden; zwei verschiedene Kurven mögen, außer eventuell T_ν, im Innern von E_ν keine gemeinsamen Punkte haben. In dem Punkt T_ν treffen sich m solche Kurven, wir nennen ihn einen *Treffpunkt* der *Koordinate* a_ν; die Kurven s_ν^μ nennen wir *Treffkurven* oder *Trefflinien* (es wird sich zeigen, daß die Kurven s_ν^μ nach der näheren Bestimmung im folgenden Absatz die Netzkurven eines Netzes von W bilden, so daß also die Treffpunkte in den algebraischen Elementargebieten die Eckpunkte des Netzes sind). Jedes Glied trennt zwei Elementargebiete, es enden auf ihm also zwei Treffkurven; beide mögen das Glied im selben inneren Punkt treffen, die Vereinigung beider nennen wir eine *Treffseite*. T_ν sei, falls E_ν ein algebraisches Elementargebiet, ein innerer Punkt von E_ν; falls E_ν ein logarithmisches Elementargebiet oder ein Strahl-Elementargebiet ist, sei T_ν der (erreichbare) Randpunkt U von E_ν (cf. 13°). Eine Treffkurve hat offenbar mit keinem Strahl des Streckenkomplexes endliche Punkte gemeinsam. Wir wählen ferner (nach 13°) die Treffkurven so, daß sie keinen endlichen Häufungspunkt haben und insbesondere also keinen endlichen Punkt eines Strahles zum Häufungspunkt haben.

Eine Treffseite s, die zwei Treffpunkte in Elementargebieten (der Koordinaten) a_i, a_k verbindet, werde durch eine topologische Funktion $f(z)$ auf denjenigen Teilbogen von L zwischen a_i, a_k abgebildet, dessen Durchlaufung dem Überschreiten des von s getroffenen Gliedes nach 10° entspricht.

Wir nennen ein von Treffkurven berandetes (eventuell erst im Unendlichen) geschlossenes Polygon ein *Elementarpolygon*, wenn keine Treffkurve in seinem Inneren liegt[16]). Jedes Elementarpolygon Q_k enthält in seinem Inneren einen Knoten K_k; es kann topologisch auf das Halbblatt H_k abgebildet werden, das von K_k repräsentiert wird, in einer Weise, daß die Singularitäten am Ende der verheftenden Teilbogen von H_k in die am Rand von Q_k liegenden Treffpunkte derselben Koordinate abgebildet werden. Die Abbildungsfunktion ist bei geeigneter Wahl eine Fortsetzung von $f(z)$ — wir bezeichnen sie ebenfalls durch $f(z)$ —, denn die Singularitäten von H_k liegen über den Koordinaten der

[16]) Elementarpolygone können den Punkt U zum mehrfachen Randpunkt haben; die Struktur eines Elementarpolygons in der Umgebung von U kann sehr kompliziert sein.

Treffpunkte am Rand von Q_k und über den Häufungspunkten dieser Koordinaten. Die topologischen Bilder der algebraischen Singularitäten am Rand von H_k sind gerade die Treffpunkte $T_\nu \neq U$ am Rand von Q_k, alle anderen Singularitäten werden in U abgebildet. Eine Treffseite s am Rand von Q_k zwischen zwei Treffpunkten enthält keinen Treffpunkt im Inneren; sie ist topologisches Bild desjenigen singularitätenfreien Teilbogens über L, über den H_k mit $H_m = f(Q_m)$ zusammenhängt, wenn K_m der Innenknoten des Elementarpolygons Q_m ist, den man von K_k aus erreicht, indem man das von s geschnittene Glied des Streckenkomplexes durchläuft: Die beiden Elementarpolygone Q_k, Q_m stoßen also genau längs derjenigen Treffseite zusammen, über deren topologisches Bild die zugeordneten Halbblätter verheftet sind. Die aus den Elementarpolygonen gebildete Figur ist also ein Netz der Fläche W. Wir haben so eine topologische Abbildung der Fläche W in die punktierte Ebene $|z| < \infty$, vermittelt durch $f^{-1}(w) = h(w)$, die Umkehrung von $f(z)$. Die Konstruktion des Netzes, die eine topologische Abbildung von W auf die punktierte Ebene darstellt, zeigt wieder die Schlichtartigkeit und den einfachen Zusammenhang von W. Sie hat außerdem noch besondere Bedeutung für die Behandlung von Randstellenproblemen im folgenden Abschnitt.

Die Randstellen

§ 13. Wie ist nun die Verteilung der Randstellen bei einer im Streckenkomplex dargestellten einfach zusammenhängenden RIEMANNschen Fläche?

Jeder Zielweg $\mathfrak{a}$ einer analytischen Funktion $w(z)$ mit dem Zielwert a definiert eine Randstelle[17]) der erzeugten RIEMANNschen Fläche W über dem Zielwert a — um lästige Fallunterscheidungen zu vermeiden, sei hier und bei den späteren Untersuchungen dieses Abschnitts stets $a \neq \infty$ angenommen. Zwei verschiedene Zielwege $\mathfrak{a}_1$, $\mathfrak{a}_2$ mit demselben Zielwert a definieren dann und nur dann dieselbe Randstelle der Fläche, wenn es möglich ist, zu jedem $\varepsilon > 0$ einen die beiden Zielwege verbindenden dritten Weg $\mathfrak{c}_\varepsilon$ zu finden, auf dem $|w(z) - a| < \varepsilon$ bleibt. $\mathfrak{c}_\varepsilon$ nennen wir eine *ε-Verbindung* oder einen *ε-Verbindungsweg* der beiden Zielwege. Die Gesamtheit der Stellen auf der Fläche, deren Urbilder durch eine ε-Verbindung mit dem Zielweg $\mathfrak{a}$ verbunden werden können, heißt eine *ε-Umgebung* U_ε der durch den Zielweg definierten Randstelle. U_ε ist eine Überlagerung der Kreisscheibe $|w - a| < \varepsilon$. Bekannt ist die folgende (unvollständige) Klassifikation[17]): Eine Randstelle heißt *unmittelbar*, wenn für hinreichend kleines $\varepsilon > 0$ „in keinem Blatt von U_ε ein innerer Punkt über $w = a$ liegt", d.h. wenn die Spurpunkte von U_ε schon in der punktierten Kreisscheibe $0 < |w - a| < \varepsilon$ liegen; eine Randstelle heißt *mittelbar*, wenn für hinreichend kleines $\varepsilon > 0$ „in jedem Blatt von U_ε ein innerer Punkt über $w = a$ liegt", d.h. für die von uns betrachteten Flächen, die nur über einer geschlossenen doppelpunktfreien JORDAN-Kurve L Singularitäten tragen: wenn jeder Weg, der von einem beliebigen nicht über L liegenden Punkt $P \in U_\varepsilon$ aus gegen die Koordinate a strebt ohne L ander-

[17]) Vgl. IVERSEN [3], ULLRICH [8].

weitig zu treffen, über a einen inneren Punkt findet. Es gibt Randstellen, die weder mittelbar noch unmittelbar sind.

Bei topologischer Abbildung des Definitionsgebietes der Funktion $w(z)$ geht ein Zielweg in einen Zielweg über, zwei Zielwege, die durch eine ε-Verbindung verbunden werden können, gehen in ebensolche über. Daher kann bei der Erklärung der Begriffe „Zielweg", „Randstelle" an Stelle der ein analytisches Bild erzeugenden Funktion $w(z)$ die ein topologisches Bild erzeugende Funktion $f(z)$ ebenso verwendet werden. Ebensowenig ändert eine topologische Abbildung der w-Ebene, die die Kurve L in eine vorgegebene Kurve L' mit denselben Eigenschaften deformiert, etwas am Zielwertcharakter eines Weges oder daran, daß zwei Wege gleiche oder verschiedene Randstellen definieren.

§ 14. Zur Untersuchung des Zusammenhanges zwischen dem Streckenkomplex S^* einer RIEMANNschen Fläche W und ihren Randstellen sei daher der Einfachheit halber angenommen, L sei die reelle Achse, orientiert in Richtung wachsenden Realteiles $\Re(w)$; das durch $f(z)$ vermittelte Bild eines die obere bzw. die untere Halbebene repräsentierenden Knotens sei der Punkt $w=i$ bzw. $w\doteq -i$; einem Glied $g(a, b)$, dessen Überschreiten dem Durchlaufen der reellen Achse in Richtung der Orientierung von a nach b entspricht, ordne $f(z)$ das durch die drei Punkte i, $l(a, b)$, $-i$ bestimmte Orthokreisbogenstück $k(a, b)$ zur reellen Achse zu, welches die Endpunkte $i, -i$ hat und die reelle Achse im Punkt $l(a, b)$ schneidet [vgl. (1a')]. Das Netz werde nach § 12 so zum Streckenkomplex S^* konstruiert, daß die Konfiguration Netz-Streckenkomplex die Bedingungen (1a), (1a'), (1b) erfüllt. Dies ist möglich, denn die Bedingung (1a'), die als einzige durch die Konstruktion nicht schon von selbst erfüllt ist, bedeutet nur, daß der Schnittpunkt der Netzseite mit einem Glied $g(a, b)$ durch $f(z)$ über die Koordinate $l(a, b)$ abgebildet wird.

Aus (1a') und aus den Eigenschaften der Funktion $l(a, b)$ erhält man nach einer kleinen Überlegung drei Lemmen.

Lemma 1. *Jeder Zielweg, der in einem logarithmischen Elementargebiet oder in einem Strahl-Elementargebiet der Koordinate a verläuft, hat den Zielwert a.*

Lemma 2. *Stoßen zwei logarithmische Elementargebiete längs einer unendlichen Gliederfolge aneinander, so ist ein Weg, der diese Gliederfolge immer wieder trifft, kein Zielweg.*

Lemma 3. *a) Jeder Weg, der in einem Strahl-Elementargebiet zum Punkt U läuft, ist Zielweg mit der Koordinate des Elementargebietes als Zielwert; b) für zwei verschiedene derartige Zielwege gibt es zu jedem $\varepsilon>0$ eine ε-Verbindung.*

§ 15. Sei gegeben ein Zielweg $\mathfrak{a}=\mathfrak{a}(t)$, $0\leq t<\infty$, im Netz. Ohne Beschränkung der Allgemeinheit dürfen wir annehmen, er sei doppelpunktfrei, treffe keinen Knoten im Netz und sein Zielwert sei Null; von einer Stelle ab trifft er entweder überhaupt kein Glied von S^* mehr, oder er trifft immer wieder Glieder. Im ersten Fall liegt er entweder auf einem Strahl — dem Zielweg entspricht also ein Strahl σ des Streckenkomplexes — oder $\mathfrak{a}$ liegt

in einem logarithmischen Elementargebiet oder einem Strahl-Elementargebiet (einschließlich des Randes) [denn algebraische Elementargebiete reichen nicht bis zum Rand des Netzes] der Koordinate Null (Lemma 1). Im zweiten Fall durchläuft $\mathfrak{a}$ von jeder Stelle an noch unendlich viele verschiedene Elementargebiete (wobei dasselbe Elementargebiet unendlich oft getroffen werden kann). [Es ist nämlich unmöglich, daß $\mathfrak{a}$ nur eine (von Eins verschiedene) endliche Anzahl n von Elementargebieten immer wieder trifft: Weil $\mathfrak{a}$ zum Rand des Netzes läuft, scheiden algebraische Elementargebiete, wegen der endlichen Anzahl Strahl-Elementargebiete aus, während man nach Lemma 2 logarithmische Elementargebiete ausschließt]. Wir dürfen annehmen, $\mathfrak{a}$ habe dann weder mit Gliedern noch mit Strahlen des Streckenkomplexes ein Kontinuum gemeinsam. Sei nun a_α die Koordinate des Elementargebietes E_α aus der von $\mathfrak{a}$ durchlaufenen Folge $\{E_\alpha\}$ von Elementargebieten. Vom Zielweg mehrfach getroffene Elementargebiete kommen in der Folge $\{E_\alpha\}$ entsprechend oft vor.

Die nach der Reihenfolge der Durchlaufung geordnete Folge $\{E_\alpha\}$ braucht nicht den Ordnungstypus ω der Folge der natürlichen Zahlen zu haben. Ordnet man jedem Elementargebiet $E_\alpha \in \{E_\alpha\}$ den Parameterwert t_α eines Punktes von $\mathfrak{a}$ innerhalb E_α zu, so erhält man eine zu $\{E_\alpha\}$ ähnliche geordnete, nirgends dichte Menge $\{t_\alpha\}$ in $0 \leq t < \infty$; jedem Häufungspunkt t' der Folge $\{t_\alpha\}$ entspricht eine Häufungsmenge $E_{t'}$ von $\{E_\alpha\}$, also ein Strahl σ des Streckenkomplexes. Die Gesamtheit $\{\bar{E}_\alpha\}$ aus den Elementargebieten E_α und ihren Häufungsmengen bildet also eine nach wachsendem Parameter t geordnete Menge; wir nennen $\alpha = t_\alpha$ die *Indexzahl* eines Elementargebietes oder Strahles in der Folge $\{\bar{E}_\alpha\}$. Die Menge $\{\alpha\}$ der Indexzahlen der Folge $\{\bar{E}_\alpha\}$ ist auf der Zahlengeraden abgeschlossen und nirgends dicht.

In $\{\bar{E}_\alpha\}$ kommen entweder unendlich viele nur logarithmische, oder endlich viele logarithmische und unendlich viele algebraische, oder unendlich viele logarithmische und unendlich viele algebraische Elementargebiete noch von jeder Stelle des Zielweges ab vor. Es sei $\{E_{\alpha_\nu}\}$ $\nu = 1, 2, \ldots$, eine abzählbar unendliche Teilfolge verschiedener Elementargebiete von $\{E_\alpha\}$ derart, daß $\lim \alpha_\nu = \infty$. Entweder gibt es in $\{E_\alpha\}$ eine Unterfolge $\{E_{\alpha_\nu}\}$, die gegen einen Strahl konvergiert, oder jede Unterfolge $\{E_{\alpha_\nu}\}$ hat als Häufungspunkt [18]) nur den Punkt U.

Jedes Elementargebiet E_α, von E_0 abgesehen, hat als Nachbarn in der Folge $\{\bar{E}_\alpha\}$ entweder zwei andere Elementargebiete, oder ein Elementargebiet und einen Strahl, oder zwei Strahlen. Es sei E_α ein Elementargebiet, das als rechten (linken), also späteren (früheren) Nachbarn in der Folge $\{\bar{E}_\alpha\}$ ein Elementargebiet E_{α^*} hat; c_α sei eine Netzseite, die die Treffpunkte T_α, T_{α^*} der Elementargebiete E_α, E_{α^*} verbindet. Es gilt dann

Lemma 4. *Ist $a = 0$ der Zielwert des Zielweges $\mathfrak{a}$, so gibt es zu jedem $\varepsilon > 0$ eine Zahl t_ε derart, daß für jedes Elementargebiet $E_\alpha \in \{E_\alpha\}$, das einen Nachbarn E_{α^*} in der Folge $\{E_\alpha\}$ hat, mit $\alpha > t_\varepsilon$ (wenigstens) eine Netzseite c_α ein topologisches Bild $f(c_\alpha)$ besitzt, das ganz über der Scheibe $|w| < \varepsilon$ liegt.*

[18]) Vgl. Fußnote [10]).

Korollar 1. *Es ist* $|a_\alpha| < \varepsilon$, $|a_{\alpha^*}| < \varepsilon$.

Korollar 2. *Für die Koordinate a' eines vom Zielweg für $t > t_\varepsilon$ getroffenen Strahls σ' oder eines Strahl-Elementargebietes E' gilt* $|a'| \leq \varepsilon$.

Korollar 3. *Ein Strahl, der Häufungsmenge der unendliche Folge $\{E_{\alpha_\nu}\}$ von Elementargebieten ist, die ein Zielweg vom Zielwert Null durchläuft, hat die Koordinate Null.*

Lemma 4 ist eine einfache Konsequenz der beiden folgenden Hilfssätze; Korollar 1 ist trivial, die Korollare 2, 3 folgen sofort aus dem Lemma zusammen mit dem Korollar 1.

Hilfssatz 1. *Ein Glied, dessen Überschreiten der Durchlaufung des Intervalles $I(a'a'')$ auf L von a' nach a'' im Sinn der Orientierung entspricht, wird durch $f(z)$ auf einen Weg über der w-Ebene abgebildet, auf dem* $|w| \geq \min\{l(a', a''), 1\}$ *ist.*

Der Hilfssatz ist eine triviale Folge aus der Definition der Abbildung $f(z)$.

Hilfssatz 2. *Sei $\mathfrak{z}$ ein Weg, der zwei verschiedenen Seiten g', g'' eines Elementargebietes E der Koordinate a verbindet, während das Elementargebiet E' der Koordinate a' durch g' von E getrennt werde, und entweder das Elementargebiet E'' der Koordinate a'' durch g'' von E getrennt werde oder g'' ein Strahl der Koordinate $a''[=a]$ sei. Sei $\mathfrak{w}$ das durch die topologische Abbildung $f(z)$ gewonnene Bild des Weges $\mathfrak{z}$. Dann ist*

$$\text{(2)} \qquad \max_{w \in \mathfrak{w}} |w| = \max_{z \in \mathfrak{z}} |f(z)| \geq \operatorname{Min}(|a|, |a'|, |a''|, 1).$$

Beweis. Es sei zunächst g'' ein Glied des Streckenkomplexes. Die Ungleichung (2) ist trivial, wenn $\mathfrak{z}$ durch den Treffpunkt von E läuft. Andernfalls unterscheiden wir zwei Fälle. a) Die Vereinigung der beiden durch g', g'' bestimmten abgeschlossenen Teilbogen von L hat den Punkt a als inneren Punkt. Wir dürfen annehmen, die Koordinaten der Elementargebiete folgen in der Reihe a', a, a'' im Sinn der Orientierung von L aufeinander. Jeder Weg $f(\mathfrak{z})$, den einen Punkt von $k(a', a)$ mit einem Punkt von $k(a, a'')$ verbindet, genügt der Relation $\max_{z \in \mathfrak{z}} |f(z)| \geq \operatorname{Min}(|a|, \operatorname{Max}(|a'|, |a''|), 1)$, also auch der Relation (2). b) Die beiden durch g', g'' bestimmten Teilbogen von L haben den Punkt a als Randpunkt. $\mathfrak{z}$ schneidet dann eine Netzseite, die ein Intervall (a, a_0) der reellen Achse bestimmt, das mit dem Intervall (a', a) keinen inneren Punkt gemein hat. Ähnlich wie oben folgt

$$\max_{z \in \mathfrak{z}} |f(z)| \geq \operatorname{Min}(|a|, \operatorname{Max}(|a'|, |a_0|), 1),$$

und hieraus (2). Läßt man das Intervall (a, a'') sich auf den Punkt a zusammenziehen, erhält man in der Grenze den Hilfssatz für den Fall, daß g'' kein Glied, sondern ein Strahl des Streckenkomplexes ist.

Aus dem Hilfssatz 2 folgt nun für drei unmittelbar aufeinanderfolgende Elemente $E_{\alpha^*}, E_\alpha, E_{\alpha^{**}}$ der Folge $\{E_\alpha\}$ mit $\alpha^* < \alpha < \alpha^{**}$, wo E_α ein Elementargebiet, $E_{\alpha^*}, E_{\alpha^{**}}$ entweder zwei Elementargebiete, oder ein Elementargebiet

und ein Strahl sind, sogleich $\lim_{\alpha\to\infty} \mathrm{Min}(|a_\alpha|, |a_{\alpha'}|, |a_{\alpha''}|) = 0$; aus Hilfssatz 1 folgt dann nach $\lim a_\alpha = \lim a_{\alpha'} = \lim a_{\alpha''} = 0$ das Lemma 4.

§ 16. Wir erhalten als Ergebnis: Nur folgenden Elementen von S^* können Randstellen von W entsprechen:

a) einem logarithmischen Elementargebiet oder einem Strahl-Elementargebiet,

b) einem Strahl,

c) einer Folge $\{\overline{E}_\alpha\}$ von Elementargebieten und Strahlen (nicht notwendig vom Ordnungstypus ω), in der die Indexzahlen eine abgeschlossene dichte Zahlmenge $0 \leq \alpha < \infty$ mit $\lim \alpha = \infty$ durchlaufen, wenn

(c1) jede Teilfolge $E_{\alpha_\nu} \in \{\overline{E}_\alpha\}$ verschiedener Elementargebiete vom Ordnungstypus ω, die Elementargebiete beliebig hoher Indexzahl enthält, gegen den Punkt U konvergiert,

(c2) zwei verschiedene Elementargebiete oder Strahlen $\overline{E}_{\alpha'}, \overline{E}_{\alpha''} \in \{\overline{E}_\alpha\}$ durch einen JORDAN-Bogen verbunden werden können, der die zwischen $E_{\alpha'}, E_{\alpha''}$ liegenden Elementargebiete und Strahlen der Folge der Reihe nach, und sonst keine weiteren, durchläuft und insbesondere den Punkt U nicht enthält,

(c3) es für je zwei in der Folge $\{\overline{E}_\alpha\}$ benachbarte Elementargebiete $E_\alpha, E_{\alpha'}$ Netzseiten e_α gibt, die die Treffpunkte der Elementargebiete verbinden, so daß die Projektionen der Bilder $f(e_\alpha)$ gegen einen Punkt konvergieren [cf. Lemma 4].

Diese drei Gruppen a), b), c) von Elementen des Streckenkomplexes nennen wir *Randelemente* des Streckenkomplexes und bezeichnen sie der Reihe nach als *Gebiets-Randelement, Strahl-Randelement, Punkt-Randelement (Gebiets-Element, Strahl-Element, Punkt-Element);* die Konvergenzkoordinate aus (c3) heiße *Koordinate* des Punkt-Randelements.

Wir sagen, ein Zielweg $\mathfrak{a}$ *definiert* ein Gebiets-Randelement, wenn er von einer Stelle ab in einem Gebiets-Randelement, also in einem logarithmischen Elementargebiet oder einem Strahl-Elementargebiet liegt; $\mathfrak{a}$ definiert ein Strahl-Randelement σ der Koordinate a, wenn $\mathfrak{a}$ entweder schließlich im Elementarpolygon des Strahl-Randelementes (cf. 7°b) läuft und den Zielwert a hat, oder wenn $\mathfrak{a}$ eine Folge $\{\overline{E}_\alpha\}$, $0 \leq \alpha < \infty$ von Elementargebieten und Strahlen durchläuft, in der eine Unterfolge $\{E_{\alpha_\nu}\}$ vom Typ ω mit $\lim \alpha_\nu = \infty$ sich gegen σ häuft. Der Zielweg $\mathfrak{a}$ definiert ein Punkt-Randelement, wenn eine Teilfolge $\{\overline{E}_{\alpha_\nu}\}$ mit $\lim \alpha_\nu = \infty$ der von $\mathfrak{a}$ durchlaufenen Folge $\{\overline{E}_\alpha\}$ von Elementargebieten und Strahlen auch Teilfolge einer ein Punkt-Randelement bestimmenden Folge $\{\overline{E}^*_{\alpha_\nu}\}$ ist.

Über den Zusammenhang zwischen den Rand-Elementen des Streckenkomplexes und Randstellen der zugeordneten RIEMANNschen Fläche gilt

Satz 2. *Jedem Randelement von S^* ist in eindeutiger Weise eine Randstelle zugeordnet.*

Der Satz besagt ein zweifaches: 1. Es gibt zu jedem Randelement einen Zielweg, der es definiert; 2. alle ein bestimmtes Randelement definierenden Zielwege definieren dieselbe Randstelle. Den Beweis des Satzes führen wir in diesen zwei Schritten.

§ 17. Jedem Randelement entspricht zunächst (wenigstens) eine Randstelle von W: Es gibt zu jedem Randelement einen dieses Element definierenden Weg, der Zielweg ist.

a) Ein Weg, der auf einer Netzseite zum Treffpunkt eines logarithmischen Elementargebietes der Koordinate a läuft, ist Zielweg mit dem Zielwert a; ebenso ist ein Weg, der in einem Strahl-Elementargebiet der Koordinate a zum Punkt U läuft, ein Zielweg mit dem Zielwert a.

b) Sei gegeben ein Strahlelement σ der Koordinate a. Zu ihm gehört (cf. 7°) ein wohlbestimmter Knoten K_a, der in einem Elementarpolygon Q_a liegt. Sei $\mathfrak{w}$ ein Weg im Halbblatt $H_a = f(Q_a)$, der gegen die Koordinate a strebt. Sein Urbild $f^{-1}(\mathfrak{w})$ strebt gegen U (cf. § 12), ist also ein Zielweg mit dem Zielwert a.

c) Gegeben sei ein Punktelement der Koordinate a, also eine Folge $\{\overline{E}_\alpha\}$ von Elementargebieten und Strahlen, die die Bedingungen (c1), (c2), (c3) erfüllt [$0 \leq \alpha < \infty$, $\lim \alpha = \infty$, $\{\alpha\}$ nirgends dicht]; $\bar{a}_\alpha$ sei die Koordinate des Elementargebietes bzw. Strahles $\overline{E}_\alpha$. Wir dürfen $a = 0$ annehmen. Sei Σ_β die Gesamtheit der Strahlen in der Teilfolge $E^{(\beta)} = \{\overline{E}_\alpha\}_{\alpha > \beta}$ von $\{\overline{E}_\alpha\}$. Nach Bedingung (c1) ist $\bigcap_\beta \Sigma_\beta$ leer.

Ist Σ_β für ein gewisses β leer, so ist die Teilfolge $E^{(\beta)}$ eine Folge $\{E_n\}$ benachbarter Elementargebiete vom Typ ω, die Koordinate von E_n sei a_n. Sind alle Elementargebiete der Folge $\{E_n\}$ algebraisch, so ist nach (c3) die Folge $\{e_n\}$ der Netzseiten, die die Treffpunkte in aufeinanderfolgenden Elementargebieten verbinden, ein Zielweg. Enthält $\{E_n\}$ logarithmische Elementargebiete, so konstruieren wir einen Zielweg $\mathfrak{z}$ in folgender Weise. Wir nehmen an, $\mathfrak{z}$ sei schon bis zu dem Punkt $P_m \in E_m$ fortgesetzt. Ist E_m ein algebraisches Elementargebiet, so sei P_m der Treffpunkt von E_m; ist E_m ein logarithmisches Elementargebiet, so sei $P_m \in E_m$ ein Punkt einer Netzseite e_{m-1}, die die Treffpunkte T_{m-1} und T_m von E_{m-1} und E_m verbindet, mit $|f(P_m) - a_m| < 2^{-m}$. Wir betrachten die Netzseite e_m, die die Treffpunkte T_m und T_{m+1} von E_m und E_{m+1} verbindet. Ist E_m ein algebraisches Elementargebiet, so werde $\mathfrak{z}$ von P_m aus auf e_m nach P_{m+1} fortgesetzt. Ist E_m ein logarithmisches Elementargebiet, so betrachten wir auf e_m einen Punkt $Q_m \in E_m$, für den $|f(Q_m) - a_m| < 2^{-m}$. P_m und Q_m werden in E_m durch eine endliche Zahl von Netzkurven voneinander getrennt. Zwei benachbarte schneiden mit dem Rand von E_m aus E_m ein Gebiet heraus, dessen Bild den in der oberen oder den in der unteren Halbebene liegenden Teil der punktierten Scheibe $0 < |w - a_m| < \varepsilon$ für hinreichend kleines $\varepsilon > 0$ enthält. Es können daher P_m und Q_m durch einen Weg $\mathfrak{c}_m$ verbunden werden, dessen Bild die 2^{-m}-Umgebung der Koordinate a_m nicht verläßt. Ist E_m ein logarithmisches Elementargebiet, so werde $\mathfrak{z}$ von P_m aus auf $\mathfrak{c}_m$ und dem zwischen Q_m und P_{m+1} liegenden Teil von e_m nach P_{m+1}

fortgesetzt. In dieser Weise fortsetzend erhält man auf Grund von (c3) einen Zielweg mit dem Zielwert $a = \lim a_m$ $(=0)$.

Ist dagegen Σ_β für kein β leer, so sei $\sigma_0 \in \Sigma_{\beta_0}$ ein beliebiger Strahl. Er definiert ein Halbblatt H_0. In H_0 gibt es nach der Ordnungseigenschaft von $\{\bar{E}_\alpha\}$ einen *letzten* Strahl $\sigma_0' \in \{\bar{E}_\alpha\}$, der insbesondere selbst höchstens einseitige Häufungsmenge anderer Strahlen, und zwar niedrigerer Indexzahlen, ist [man beachte, daß jeder Strahl in einem wohlbestimmten Elementarpolygon liegt]. Es gibt daher zwei andere Indexzahlen β_0', β_1 mit $\beta_0 < \beta_0' \leq \beta_1$ derart, daß $\bigcup_{\beta_0'}^{\beta_1} \bar{E}_\alpha$ keinen Strahl enthält und daß ebenso zwischen σ_0' und $E_{\beta_0'}$ in der Folge $\{\bar{E}_\alpha\}$ kein Strahl vorkommt. In $\Sigma_{\beta_0'}$ gibt es wieder einen ersten Strahl σ_1. Dieser definiert ein Halbblatt H_1, in diesem gibt es wieder einen letzten Strahl σ_1', usf. Auf diese Weise erhält man eine Folge H_ν von lauter verschiedenen Halbblättern und von ersten und letzten Strahlen σ_ν, σ_ν' in ihnen [$\sigma_\nu = \sigma_\nu'$ ist natürlich möglich]. Diese Folge kann nicht abbrechen, weil Σ_β für kein β leer ist, überdies haben die Strahlen σ_ν, σ_ν' in $\{\bar{E}_\alpha\}$ für hinreichend großes ν beliebig große Indexzahlen α [cf. (c2)]: Denn ein Weg, der sämtliche Strahlen σ_ν, σ_ν' $(\nu = 1, 2, \ldots)$ trifft, konvergiert notwendig gegen den Punkt U, weil eine unendliche Menge von Strahlen, von denen höchstens zwei denselben Knoten bestimmen, außer dem Punkt U keinen Häufungspunkt hat [cf. § 12, Abs. 1, und (c 2)].

Ein Zielweg, der das gegebene Punkt-Randelement $\{\bar{E}_\alpha\}$ der Koordinate 0 definiert, wird auf folgende Weise konstruiert. Zunächst streben die Koordinaten m_ν, m_ν' der Strahlen σ_ν, σ_ν' gegen Null [cf. (c3)]. Auf jedem Strahl σ_ν bzw. σ_ν' wird ein Punkt P_ν bzw. P_ν' markiert, für den $|f(P_\nu) - m_\nu| < 2^{-\nu}$ bzw. $|f(P_\nu') - m_\nu'| < 2^{-\nu}$. Von P_ν, P_ν' aus wird ein so kurzer JORDAN-Bogen $\mathfrak{z}_\nu$ bzw. $\mathfrak{z}_\nu'$ nach der Seite fallender bzw. wachsender Indexzahlen in den getroffenen Elementargebieten derart gezogen, daß nur Elementargebiete der Folge $\{\bar{E}_\alpha\}$ [cf. (c2)] getroffen werden, insbesondere also kein weiterer Strahl, und daß ferner das Bild $f(\mathfrak{z}_\nu)$, $f(\mathfrak{z}_\nu')$ im Kreis $|w - m_\nu| < 2^{-\nu}$ bzw. $|w - m_\nu'| < 2^{-\nu}$ enthalten ist; letzteres ist wegen der Stetigkeit von $f(z)$ möglich. Es ende $\mathfrak{z}_\nu$ bzw. $\mathfrak{z}_\nu'$ im Punkt $B_\nu \in E_{\beta_\nu}$ bzw. im Punkt $B_\nu' \in E_{\beta_\nu'}$ und es sei $\beta_\nu' \leq \beta_{\nu+1}$. Man kann nun leicht einen Zielweg finden, der die Punkte $P_\nu, P_\nu', B_\nu', B_{\nu+1}, P_{\nu+1}$, $(\nu = 1, 2, \ldots)$ verbindet: Man verbinde dazu P_ν mit P_ν' durch einen Weg $\mathfrak{p}_\nu$, dessen Bild ganz im Kreis $|w| < |m_\nu| + |m_\nu'| + 2^{-\nu}$ liegt; das ist möglich, denn $f(P_\nu)$, $f(P_\nu')$ sind zwei Punkte in H_ν, die schon diese Bedingung erfüllen. Die Endpunkte B_ν', $B_{\nu+1}$ der JORDAN-Bögen $\mathfrak{z}_\nu'$, $\mathfrak{z}_{\nu+1}$ liegen in Elementargebieten $E_{\beta_\nu'}$, $E_{\beta_{\nu+1}}$, die entweder zusammenfallen oder in der Folge $\{\bar{E}_\alpha\}$ durch endlich viele sich aneinanderschließende logarithmische und algebraische Elementargebiete getrennt sind, das Maximum der Beträge ihrer Koordinaten sei $m_{\nu,\nu+1}$. Im letzten Fall, oder wenn im ersten Fall $E_{\beta_\nu'} = E_{\beta_{\nu+1}}$ kein Strahl-Elementargebiet ist, können B_ν', $B_{\nu+1}$ durch einen Weg $\mathfrak{p}_{\nu,\nu+1}$ verbunden werden, dessen Bild in $|w| \leq |m_\nu| + |m_\nu'| + m_{\nu,\nu+1} + 2^{-\nu} + k_\nu$ $(\lim k_\nu = 0)$ enthalten ist [cf. (c2), (c3) und die für $\cap \Sigma_\beta = \emptyset$ besprochene Konstruktion]. Ist

aber $E_{\beta'_\nu} = E_{\beta_{\nu+1}}$ ein Strahl-Elementargebiet, so können B'_ν, $B_{\nu+1}$ — wie eine kleine Überlegung unmittelbar zeigt — durch einen Weg $\mathfrak{p}_{\nu,\nu+1}$ mit derselben oben angegebenen Eigenschaft verbunden werden. Da nach (c3) auch $\lim m_{\nu,\nu+1} = 0$, ist der Weg $\overset{\infty}{\underset{\nu=0}{\cup}} \{\mathfrak{z}_\nu \cup \mathfrak{p}_\nu \cup \mathfrak{z}'_\nu \cup \mathfrak{p}_{\nu,\nu+1}\}$ ein Zielweg mit dem Zielwert Null, der das Punkt-Randelement $\{\bar{E}_\alpha\}$ definiert.

§ 18. Allen drei Arten von Randelementen ist aber auch *nur eine* Randstelle von W zugeordnet. Für Strahl-Elementargebiete ist dies bereits in Lemma 3 ausgesagt. Für die übrigen Randelemente erkennen wir das auf folgende Weise. Zwei verschiedene Zielwege, die dasselbe Randelement definieren, haben (cf. Lemma 1 für logarithmische Elementargebiete, Lemma 4 mit Korollaren für Strahl- und Punkt-Randelemente) jedenfalls denselben Zielwert.

Seien zunächst in einem logarithmischen Elementargebiet E der Koordinate $a \neq \infty$ zwei Zielwege $\mathfrak{a}$, $\mathfrak{a}'$ gegeben. $\mathfrak{a}$ sei, wie wir ohne Beschränkung der Allgemeinheit annehmen dürfen, eine Netzkurve. Eine ähnliche Überlegung, wie sie in § 17c) Abs. 2 angestellt wurde, um zu zeigen, daß die Punkte P_m, Q_m in einem logarithmischen Elementargebiet E_m durch eine Kurve verbunden werden können, auf deren Bild $|w - a_m| < 2^{-m}$ bleibt, ergibt, daß jeder Punkt $P \in \mathfrak{a}'$ mit $|f(P) - a| < \varepsilon$ für hinreichend kleines $\varepsilon > 0$ durch eine im Elementargebiet E verlaufende ε-Verbindung $\mathfrak{c}_\varepsilon$ mit dem Weg $\mathfrak{a}$ verbunden werden kann; beide Zielwege definieren also dieselbe Randstelle. Die Randstelle ist genau dann ein *logarithmischer Windungspunkt*, wenn die Koordinaten der Nachbargebiete von E im Streckenkomplex die Zahl a nicht zum Häufungspunkt haben. Dann und nur dann ist nämlich für hinreichend kleines $\varepsilon > 0$ das Netzbild der ε-Umgebung der Randstelle, die durch E bestimmt ist, ganz in E enthalten und — mit $a = 0$ — ein Zweig von $\log w$ eine in der ε-Umgebung der Randstelle eindeutige und einwertige Funktion.

Ein Strahl-Randelement σ der Koordinate $a \neq \infty$ definiert zunächst eindeutig den Knoten K, von dem aus es nach U läuft. Sei $\mathfrak{a}$ ein σ definierender Zielweg, der in dem Elementarpolygon Q verläuft, das K enthält, und sei ein weiterer Zielweg $\mathfrak{a}'$ gegeben. Entweder verläuft $\mathfrak{a}'$ schließlich ebenfalls in Q — dann definiert $\mathfrak{a}'$ trivialerweise dieselbe Randstelle wie $\mathfrak{a}$ — oder $\mathfrak{a}'$ durchläuft (unter anderen) eine Folge $\{E_n^*\}$ von logarithmischen und algebraischen Elementargebieten der Koordinaten a_n^*, die alle K als Eckknoten haben und gegen σ konvergieren, während von jeder Stelle ab $\mathfrak{a}'$ noch unendlich viele Elementargebiete der Folge $\{E_n^*\}$ trifft. Sei $\varepsilon > 0$ gegeben. Wir dürfen annehmen, m sei so groß, daß für die Koordinate a_m^* von $E_m^* \in \{E_n^*\}$ die Ungleichung $|a_m^* - a| < \frac{\varepsilon}{4}$ gilt, und daß für einen Punkt $P_m \in E_m^*$ von $\mathfrak{a}'$ gilt $|f(P_m) - a| < \frac{\varepsilon}{4}$, also $|f(P_m) - a_m^*| < \frac{\varepsilon}{2}$. Ist E_m^* ein algebraisches Elementargebiet, zu dem der algebraische Windungspunkt A_m^* auf W gehört, so ist auf der Geraden $\mathfrak{g}_m = \overline{f(P_m)\, A_m^*}$ offenbar $|w - a| \leq |w - a_m^*| + |a_m^* - a| < \varepsilon$. Der am Rand von

$f(Q)$ liegende Punkt A_m^* kann offenbar mit $f(\mathfrak{a})$ in $f(Q)$ durch einen Weg $\mathfrak{w}_m$ verbunden werden, dessen Projektion sich von der Koordinate a um weniger als ε entfernt. Das Urbild von $\mathfrak{w}_m \cup \mathfrak{g}_m$ ist dann eine ε-Verbindung beider Zielwege. Ist E_m^* ein logarithmisches Elementargebiet, so kann (cf. § 17) P_m mit einem Punkt $Q_m \in E_m^*$ am Rand von Q durch eine Kurve verbunden werden, auf der $|w - a_m^*| < \frac{\varepsilon}{2}$, also $|w - a| < \varepsilon$ bleibt. Zwischen Q_m und $\mathfrak{a}$ kann wie oben eine ε-Verbindung gefunden werden; die Zielwege $\mathfrak{a}$, $\mathfrak{a}'$ definieren also dieselbe Randstelle. Ganz ähnlich sieht man ein, daß zwei Zielwege, die dasselbe Punkt-Randelement definieren, auch dieselbe Randstelle definieren, denn beide haben von jeder Stelle ab immer wieder Elementargebiete gemein, in deren Bildern sich die Bilder der Zielwege schließlich beliebig nahekommen. Satz 2 ist somit bewiesen.

§ 19. Bisher haben wir gefunden, daß jedem Randelement des Streckenkomplexes in eindeutiger Weise eine Randstelle der dargestellten RIEMANNschen Fläche entspricht. Um den Zusammenhang zwischen Randelementen und Randstellen völlig zu klären, muß folgende Frage beantwortet werden: *Wann gehört zu zwei verschiedenen Randelementen R_1, R_2 dieselbe Randstelle der Fläche?* [Beide Randelemente müssen natürlich dieselbe Koordinate haben.] Auskunft hierüber gibt

Satz 3. *Notwendig und hinreichend dafür, daß zwei Randelemente R_1, R_2 der Koordinate $a \neq \infty$ dieselbe Randstelle über a bestimmen, ist, daß sie für jedes $\varepsilon > 0$ durch eine Folge $\{\overline{E}_\alpha(\varepsilon)\}$ von Elementargebieten und Strahlen verbunden sind, die folgende Bedingungen erfüllen.*

(B_1) *Für jedes z auf einer (geeigneten) Netzseite, die die Treffpunkte von algebraischen oder logarithmischen Elementargebieten verbinden, die Nachbarn in der Folge $\{\overline{E}_\alpha(\varepsilon)\}$ sind, ist $|f(z) - a| < \varepsilon$.*

(B_2) *Die Koordinaten $\bar{a}_\alpha(\varepsilon)$ der $\bar{E}$lemente der Folge $\{\overline{E}_\alpha(\varepsilon)\}$ unterscheiden sich dem Betrag nach um höchstens ε von der Zahl a.*

Die Randelemente R_1, R_2 heißen *verbunden durch* $\{\overline{E}_\alpha(\varepsilon)\}$, wenn

(V_1) die Indexzahlen α eine abgeschlossene nirgends dichte Zahlmenge $0 \leq \alpha \leq 1$ bilden;

(V_2) das erste Element $\overline{E}_0(\varepsilon)$ der Folge $\{\overline{E}_\alpha(\varepsilon)\}$ ein Elementargebiet des Randelementes R_1 oder, falls R_1 ein Strahl σ_1 ist, dieser Strahl ist, das letzte Element $\overline{E}_1(\varepsilon)$ der Folge $\{\overline{E}_\alpha(\varepsilon)\}$ ein Elementargebiet des Randelementes R_2 oder, falls R_2 ein Strahl σ_2 ist, dieser Strahl ist;

(V_3) die Folge $\{\overline{E}_\alpha(\varepsilon)\}$ die Bedingung (c2) aus § 16 erfüllt.

Ein Elementargebiet eines Gebiets-Randelementes ist hierbei nur das logarithmische Elementargebiet bzw. das Strahl-Elementargebiet selbst, ein Elementargebiet eines Strahl- oder Punkt-Randelementes ist ein Elementargebiet einer Folge $\{\overline{E}_\alpha\}$, die das Randelement definiert. Strahl-Randelemente, die vom selben Knoten ausgehen und dieselbe Koordinate haben, bestimmen

danach[19]) — wie auch leicht direkt zu ersehen — dieselbe Randstelle. Aus Satz 3 folgt sofort [cf. Korollare 1, 2 von Lemma 4] als

Korollar. *Notwendig dafür, daß R_1, R_2 dieselbe Randstelle definieren, ist, daß R_1, R_2 sich für jedes $\varepsilon > 0$ durch eine Folge $\{\overline{E}_\alpha(\varepsilon)\}$ von Elementargebieten und Strahlen verbinden lassen, für deren Koordinaten $\bar{a}_\alpha(\varepsilon)$ die Beziehung $|\bar{a}_\alpha(\varepsilon) - a| < \varepsilon$ gilt.*

Beweis. Sind die zwei Randelemente R_1, R_2 ein logarithmisches Elementargebiet E_l und ein Strahl σ, der Seite von E_l ist, bzw. zwei Strahl-Elemente σ_1, σ_2, die beide Seiten eines Strahl-Elementargebietes E_s sind, dann definieren offenbar R_1, R_2 einerseits dieselbe Randstelle, andererseits erfüllen sie die Bedingungen $(B_1), (B_2)$ [es bilden $\{\sigma, E_l\} = \{\overline{E}_0, \overline{E}_1\}$ bzw. $\{\sigma_1, E_s, \sigma_2\} = \{\overline{E}_0, \overline{E}_{\frac{1}{2}}, \overline{E}_1\}$ für jedes $\varepsilon > 0$ eine Folge $\{\overline{E}_\alpha(\varepsilon)\}$]; stehen R_1, R_2 in dieser speziellen Beziehung, gilt also Satz 3. Daß aber auch im allgemeinen die Bedingungen (B_1), (B_2) zunächst hinreichend sind, ist trivial nach den Konstruktionen des § 17, denn ähnliche Konstruktionen, wie diejenigen, die im § 17 einen ein Randelement definierenden Zielweg lieferten, liefern eine ε-Verbindung der beiden Zielwege, die R_1, R_2 definieren. Die Notwendigkeit von $(B_1), (B_2)$ erkennt man aus folgender Betrachtung, in der ohne Beschränkung der Allgemeinheit $a \neq 0$ angenommen sei.

Angenommen, die Randelemente R_1, R_2 definieren dieselbe Randstelle, während es für hinreichend kleines $\varepsilon > 0$ nicht möglich sei, R_1 und R_2 durch eine Folge $\{\overline{E}_\alpha(\varepsilon)\}$ mit den Eigenschaften $(B_1), (B_2)$ zu verbinden. Zunächst einmal ist die Bedingung (B_1) nur dann leer, wenn R_1, R_2 entweder ein Strahl σ und ein logarithmisches oder Strahl-Elementargebiet mit der Seite σ sind, oder wenn R_1, R_2 zwei Strahlen sind, die ein Strahl-Elementargebiet beranden, also in dem bereits erledigten Fall. In allen anderen Fällen ist die Bedingung (B_1) nicht leer, wir dürfen unter Benutzung von Hilfsatz 1 also annehmen, es gebe in jeder verbindenden Folge $\{\overline{E}_\alpha(\varepsilon)\}$ [also einer Folge, die $(V_1), (V_2), (V_3)$ erfüllt] wenigstens zwei in der Folge benachbarte Elementargebiete E', E'' mit der Eigenschaft

$(\overline{B})$ *jede Netzseite e', die die Treffpunkte T', T'' von E', E'' verbindet, hat ein Bild $f(e')$, das die ganze Strecke $M_1 \leq \Re(w) \leq M_2 (M_1 < M_2)$ der reellen Achse überdeckt.*

Die Annahme, es gebe dennoch eine ε-Verbindung der beiden Zielwege $\mathfrak{a}_1, \mathfrak{a}_2$, die die Randelemente R_1, R_2 definieren, führt auf einen Widerspruch: Sei nämlich $\mathfrak{c}_\varepsilon$ eine ε-Verbindung der beiden Zielwege, die die Elementargebietsfolge $\{\overline{E}_\alpha(\varepsilon)\}$ definiert. Weil für $z \in \mathfrak{c}_\varepsilon$, insbesondere für den Schnittpunkt von $\mathfrak{c}_\varepsilon$ mit dem E', E'' trennenden Glied e', auch $|f(z)| < \varepsilon$ und weil $\varepsilon > 0$ beliebig klein angenommen werden kann, muß mit $(\overline{B})$ nach Hilfssatz 1 dann $M_1 < 0 < M_2$ gelten. Über $(\overline{B})$ hinaus haben wir daher mit

$$M = \mathrm{Min}(-M_1, M_2):$$

$(\overline{B}^*)$ *$f(e')$ überdeckt die ganze Strecke $|\Re(w)| < M$, $M > 0$, der reellen Achse.*

[19]) Vgl. § 7, vorletzter Absatz, und Lemma 3.

Es ist insbesondere also $|a'| \geq M$, $|a''| \geq M$ mit a', a'' als Koordinaten von E', E''. Entweder hat wenigstens eines der Elementargebiete E', $E'' \in \{\overline{E}_\alpha(\varepsilon)\}$ zwei Elementargebiete als Nachbarn in der Folge $\{\overline{E}_\alpha\}$ (die auch übereinstimmen können), oder es sind E', E'' benachbarte logarithmische Elementargebiete, die beide je einen Strahl σ' bzw. σ'' am Rand besitzen, wobei jeder dieser Strahlen von $\mathfrak{c}_\varepsilon$ getroffen wird. Im ersten Fall habe etwa E' außer E'' noch ein Elementargebiet E der Koordinate a als Nachbarn in der Folge. Nach Hilfssatz 2 gilt für die Koordinate a von E die Ungleichung $|a| < \varepsilon$. Wegen $l(-M, 0) < 0$ und $l(0, M) > 0$ und der Stetigkeit und Monotonie der Funktion $l(a, b)$ in beiden Argumenten gibt es ein (kleines) positives ε_1, so daß $l(-M, u) < \gamma_1 < 0$ und $l(u, M) > \gamma_2 > 0$ für $|u| < \varepsilon_1$. Für den Schnittpunkt P von $\mathfrak{c}_\varepsilon$ mit dem E, E' trennenden Glied gilt also für $\varepsilon \leq \varepsilon_1$ die Beziehung $|f(P)| > \mathrm{Min}(-\gamma_1, \gamma_2) = \gamma > 0$. Es gibt also auf dem Weg $\mathfrak{c}_\varepsilon$ Stellen, an denen $|f(z)| > \gamma > 0$. Wählt man $0 < \varepsilon < \mathrm{Min}(\varepsilon_1, \gamma)$, so ist im ersten Fall $\mathfrak{c}_\varepsilon$ gegen die Annahme keine ε-Verbindung der Zielwege $\mathfrak{a}_1$, $\mathfrak{a}_2$. Im zweiten Fall folgt ein Widerspruch daraus, daß für $z \in \sigma'$ bzw. $z \in \sigma''$ auch $|f(z)| > \mathrm{Min}(|a'|, |a''|, 1) > \varepsilon$ ist. Beide Zielwege definieren also verschiedene Randstellen über $w = 0$. Damit ist der Beweis von Satz 3 erbracht.

Beispiele

§ 20. Drei Beispiele mögen nun die bisherigen Ausführungen veranschaulichen.

a) $w_1(z) = z \sin z$ erzeugt eine RIEMANNsche Fläche W_1. Die algebraischen Singularitäten liegen über der reellen Achse, ihre Koordinaten sind die Funktionswerte von $w_1(z)$ in den (reellen) Nullstellen der Ableitung. Transzendente Singularitäten liegen nur über der Koordinate $w = \infty$, dem einzigen Zielwert von $w_1(z)$. Zur Konstruktion des Streckenkomplexes $S^*(W_1)$ kann als doppelpunktfreie geschlossene Zerschneidungskurve L die reelle Achse gewählt werden. Durch Bestimmung der reellen Züge von $w_1(z)$ erhält man ein Netz und hieraus den Streckenkomplex $S^*(W_1)$ der Fig. 1. $S^*(W_1)$ besteht aus einer Leiter von viereckigen algebraischen Elementargebieten, die nach beiden Seiten ins Unendliche führt; er ist bis auf die Koordinaten der algebraischen Windungspunkte gleich dem Streckenkomplex S der von $\sin z$ erzeugten Fläche, dem Sinus-Streckenkomplex, und hat, wie dieser, zwei logarithmische Elementargebiete der Koordinate ∞. $S^*(W_1)$ ist (einschließlich der Koordinaten der Elementargebiete) symmetrisch zum algebraischen Elementargebiet E_0 der Koordinate $a_0 = 0$. Das in der Leiter an n-ter Stelle auf E_0 folgende Elementargebiet sei E_n, seine Koordinate a_n hat das Vorzeichen von $(-1)^{n-1}$. Es wächst $(-1)^{n-1} a_n \sim n$ monoton gegen ∞. Nach Satz 3 bestimmen die beiden logarithmischen Elementargebiete der Koordinate ∞ nur eine einzige Randstelle über ∞, sie ist trivialerweise unmittelbar, weil W_1 über $w_1 = \infty$ überhaupt keinen (inneren) Punkt hat; sie ist aber kein logarithmischer Windungspunkt [§ 18, Abs. 2]. Die beiden in Fig. 1 nach links und rechts führenden Folgen viereckiger Elementargebiete bestimmen nach Satz 2

keine Randstelle, denn sie definieren kein Rand-Element: Die Bilder der Netzseiten, die zwei aufeinanderfolgende Treffpunkte verbinden, konvergieren nicht gegen einen Punkt (sondern sie streben sogar gegen die in $w_1 = \infty$ punktierte reelle Achse). Jede andere Folge von Elementargebieten, die ein Randelement von $S^*(W_1)$ definiert, enthält immer wieder eines der logarithmischen Elementargebiete, definiert also dieselbe Randstelle wie dieses. W_1 hat an Randstellen also nur eine einzige, und zwar eine unmittelbare, über $w_1 = \infty$.

b) $w_2(z) = z^{-1}\sin z$ erzeugt eine RIEMANNsche Fläche W_2. Wie bei W_1 liegen die algebraischen Singularitäten von W_2 nur über der reellen Achse, transzendente liegen nur über den Koordinaten $w_2 = 0, \infty$ den einzigen Zielwerten von $w_2(z)$. Wie oben erhält man den Streckenkomplex $S^*(W_2)$ in Fig. 2. $S^*(W_2)$ ist wieder bis auf die Koordinaten der algebraischen Elementargebiete der Sinus-Streckenkomplex. Wie $S^*(W_1)$ ist $S^*(W_2)$ symmetrisch

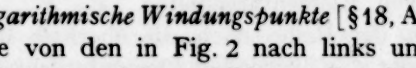


Fig. 1

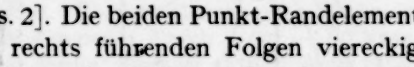


Fig. 2

zu einem Elementargebiet E_0; dessen Koordinate a_0 ist Eins. Für die Koordinate a_n des auf E_0 in der Leiter an n-ter Stelle folgende Elementargebiet gilt $(-1)^n \cdot a_n \sim n^{-1}$ monoton fallend. Die beiden von den logarithmischen Elementargebieten gebildeten Randelemente der Koordinate ∞ bestimmen (nach dem Korollar von Satz 3) zwei *verschiedene* Randstellen über $w_2 = \infty$, und zwar *logarithmische Windungspunkte* [§18, Abs. 2]. Die beiden Punkt-Randelemente, die von den in Fig. 2 nach links und rechts führenden Folgen viereckiger Elementargebiete gebildet werden, bestimmen (nach Satz 2) Randstellen über $w_2 = 0$, und zwar (nach Satz 3) zwei verschiedene; diese sind trivialerweise mittelbar, denn von jedem Punkt $P \in W_2$ gelangt man bei radialer Annäherung an die Koordinate Null zu einem Flächenpunkt über $w_2 = 0$, weil in keinem Halbblatt von W_2 über $w_2 = 0$ eine Singularität liegt. $S^*(W_2)$ hat offenbar keine weiteren Randelemente, W_2 also keine weiteren Randstellen.

c) $w(z) = e^z - z^{-1}$ erzeugt eine RIEMANNsche Fläche W_3. Die unendlich vielen Nullstellen von $w'(z) = e^z + z^{-2}$ sind alle einfach und nicht-reell; sie liegen sämtlich auf der durch $|e^z| = z^{-2}$ gegebenen Kurve C. Die in $\Im(z) > 0$ liegenden seien $z_1, z_2, \ldots$, der Reihe nach angeordnet, wie sie beim Durchlaufen des in $\Im(z) > 0$ liegenden Teiles C^+ von C, ausgehend vom Schnittpunkt mit der reellen Achse, getroffen werden. Es ist

$$w_\nu = w(z_\nu) = e^{z_\nu} - z_\nu^{-1} = -z_\nu^{-2} - z_\nu^{-1}.$$

Man berechnet leicht, daß sowohl der Realteil $\Re(w_\nu)$ als auch der Imaginärteil $\Im(w_\nu)$ eigentlich monoton in ν gegen Null fallen. Die transzendenten Singularitäten von W_3 liegen über den Zielwertkoordinaten $w = 0, \infty$ von $w(z)$. Durchläuft man C^+ in der nach $z = \infty$ weisenden Richtung, so ist $\Re(z)$ monoton fallend; auf dem in $\Im(z) > 0$, $\Re(z) < 0$ liegenden Teil von C^+ ist $\Re\{g(z)\}$,

mit $g(z) = -z^{-2} - z^{-1}$; monoton fallend. Ferner ist $\Re(z_1) > 0$, $\Re(z_2) < 0$; der zwischen z_2 und ∞ liegende Teil von C^+ wird durch $g(z)$ auf einen Kurvenbogen L_1 abgebildet. Sei L_2 das w_1, w_2 verbindende Geradenstück (das mit L_1 nur den Punkt w_2 gemeinsam hat), L_3 das Geradenstück $\Re(w) = \Re(w_1)$, $\Im(w_1) \leq \Im(w) < \infty$ und seien $\bar{L}_1, \bar{L}_2, \bar{L}_3$ die Spiegelungen von L_1, L_2, L_3 an der reellen Achse. Es ist dann

$$L = L_1 \cup L_2 \cup L_3 \cup \bar{L}_1 \cup \bar{L}_2 \cup \bar{L}_3$$

eine (stückweise analytische) doppelpunktfreie geschlossene Kurve, die durch alle Koordinaten von Singularitäten der Fläche W_3 hindurchgeht. Man erhält das Netz von W_3 in den Urbildern von L in der z-Ebene und aus ihm den Streckenkomplex $S^*(W_3)$ in Fig. 3. $S^*(W_3)$ besteht aus einer Folge F von Knoten und Gliedern, die ein logarithmisches Elementargebiet E der Koordinate ∞ begrenzt; ein Innenknoten F_0 der Folge ist außerdem über ein Glied s' mit einem Außenknoten K', dieser über ein Glied s'' mit einem Innenknoten K'' verbunden; K'' selbst ist mit allen Außenknoten der Folge F durch Glieder verbunden, er ist also Eckknoten einer unendlichen Folge $\{E_\nu\}$ ($\nu = \ldots, -2, -1, 1, 2, \ldots$) von viereckigen Elementargebieten. $E_1, E_2, \ldots$ seien der Reihe nach die oberhalb der Glieder s', s'' liegenden Elementargebiete, $E_{-1}, E_{-2}, \ldots$ die unterhalb liegenden. Die Koordinate von E_ν ist w_ν für $\nu > 0$, $\overline{w}_\nu$ für $\nu < 0$. Es ist $\lim w_\nu = \lim \overline{w}_\nu = 0$.

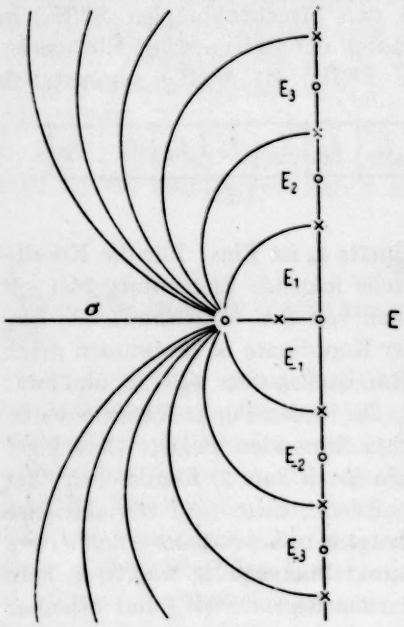


Fig. 3

$S^*(W_3)$ hat als Randelemente das logarithmische Elementargebiet E der Koordinate ∞ und ein Strahl-Randelement σ der Koordinate 0, das durch die Folge $\{E_\nu\}$ definiert wird. Ihnen entsprechen: eine Randstelle von W_3 über ∞ — die nach § 18 ein logarithmischer Windungspunkt, also eine unmittelbare Randstelle ist — und eine Randstelle über 0. Diese ist weder mittelbar noch unmittelbar: im Halbblatt H'', das zum Knoten K'' gehört, erreicht man von keinem Punkt aus bei Annäherung an die Koordinate 0 einen Flächenpunkt über 0; in allen anderen Halbblättern von W_3 hingegen liegen innere Punkte über der Koordinate 0.

Beziehungen zwischen den Streckenkomplexbegriffen

§ 21. Ist einer RIEMANNschen Fläche W ein Streckenkomplex S bzw. ein Streckenkomplex S^* zugeordnet, so wurde dieser mit $S(W)$ bzw. $S^*(W)$ bezeichnet. Trivialerweise umfaßt der Definitionsbereich $\mathfrak{S}^*$ des Funktionals S^*

den Definitionsbereich $\mathfrak{S}$ des Funktionals S. Ist $W \in \mathfrak{S}$, liegen also die Singularitäten von W über endlich vielen Grundpunkten, so erhält man $S^*(W)$ aus $S(W)$ dadurch, daß man in $S(W)$ alle zwischen zwei Knoten verlaufenden Bündel benachbarter Glieder durch ein einziges Glied ersetzt. Umgekehrt erhält man bei einer Fläche $W \in \mathfrak{S}$, deren Streckenkomplex $S^*(W)$ gegeben ist, $S(W)$ dadurch, daß man ein Glied, dessen Überschreiten dem Durchlaufen des Bogens s von L entspricht, durch ein Bündel von $k+1$ benachbarten Gliedern ersetzt, wenn k die Anzahl der Grundpunkte im Innern von s ist, über denen W überhaupt verzweigt ist[20]).

Erweiterungen in ähnlicher Richtung, wie sie in § 2 für den Streckenkomplex S erwähnt wurden, sind auch für S^* durchzuführen: Gegeben seien n einfach zusammenhängende Gebiete G_ν $(\nu = 1, 2, \ldots, n)$ in der Ebene, von denen jedes durch L in zwei Teilgebiete zerlegt wird, während zwei verschiedene Gebiete G_ν, G_μ höchstens Randpunkte gemeinsam haben. Es sei $W \in \mathfrak{S}^*$ eine Fläche, die in jedem Blatt $B_\varkappa$ höchstens je eine Singularität $a_\nu^{(\varkappa)} \in G_\nu$ enthält. Jeder Fläche, die aus W dadurch hervorgeht, daß die Windungspunkte $a_\nu^{(\varkappa)}$ innerhalb G verschoben werden, werde als Streckenkomplex $S^*(W)$ zugeordnet. Auf Einzelheiten will ich hier jedoch nicht eingehen.

Literatur

[1] ELFVING, G.: Über eine Klasse von Riemannschen Flächen und ihre Uniformisierung. Acta Soc. Sci. Fennicae, Nova Ser. **2**, Nr. 3, 1—60 (1934). — [2] HANCK, H.: Über die Ableitungsfestigkeit gewisser Verzweigungseigenschaften. Dissertation Göttingen **1935**. — [3] IVERSEN, F.: Recherches sur les fonctions inverses des fonctions méromorphes. Thèse de Helsingfors **1914**. — [4] NEVANLINNA, R.: Über die Riemannsche Fläche einer analytischen Funktion. Verh. Internat. Math. Kongreß 1 Zürich 1932. — [5] NEVANLINNA, R.: Eindeutige Analytische Funktionen. Berlin: Springer 1936; 2. Aufl. 1953. — [6] NEVANLINNA, R.: Uniformisierung. Berlin: Springer 1953. — [7] TEICHMÜLLER, O.: Vermutungen und Sätze über die Wertverteilung gebrochener Funktionen endlicher Ordnung. Dtsch. Math. **4**, 163—190 (1939). — [8] ULLRICH, E.: Flächenbau und Wachstumsordnung bei gebrochenen Funktionen. Jber. Deutsch. Math. Verein. **46**, 232—274 (1936).

[20]) Man könnte natürlich auch beim Streckenkomplex S^* (wie man es beim Streckenkomplex S tut) Mehrfachbindungen zulassen, in der Weise, daß zweiseitigen Elementargebieten nicht-singuläre Stellen der Fläche über L zugeordnet sind.

Gießen, Mathematisches Institut der Universität

(Eingegangen am 3. September 1955)

HUCKEMANN, F.
Math. Zeitschr. Bd. 65, S. 240—282 (1956)

Über den Einfluß von Randstellen Riemannscher Flächen auf die Wertverteilung*)

Herrn ROLF NEVANLINNA zum 60. Geburtstag gewidmet

Von

FRIEDRICH HUCKEMANN

Einleitung

Eine der wesentlichen Problemstellungen der modernen Funktionentheorie ist die Frage nach den gegenseitigen Beziehungen zwischen einer analytischen Funktion und der RIEMANNschen Fläche, die sie erzeugt. Insbesondere geht man von der RIEMANNschen Fläche als *primärer Gegebenheit* aus und fragt, wie ihr Verzweigungscharakter in der Wertverteilung der erzeugenden Funktion zum Ausdruck kommt. Es liegt in der Natur der Sache, daß die Randstellen der Fläche hier von entscheidendem Einfluß sind.

Die außerordentlich komplizierte Struktur von ganz allgemeinen RIEMANNschen Flächen brachte es mit sich, daß man sich zunächst nur mit einfachen Flächentypen beschäftigte. Insbesondere kam man dazu, die Klasse $\mathfrak{R}$ einfach zusammenhängender nicht-kompakter Flächen, die nur über endlich vielen Grundpunkten Windungspunkte tragen, zu studieren; dies liegt auch nahe, denn diese Flächen gestatten im Streckenkomplex eine übersichtliche Darstellung, was Untersuchungen natürlich bedeutend erleichtert, andererseits sind die Flächen doch so allgemein, daß eine Kenntnis der hier herrschenden Verhältnisse Fingerzeige und Anhaltspunkte dafür gibt, welche Aussagen man überhaupt erwarten darf über den Zusammenhang zwischen der geometrischen Struktur der Fläche und dem analytischen Verhalten der erzeugenden Funktion.

Die Flächen aus $\mathfrak{R}$ besitzen als Randstellen allerdings nur logarithmische Windungspunkte, eine gewiß nicht sehr allgemeine Art von Randstellen; es läßt sich jedoch eine wesentlich umfassendere Klasse $\mathfrak{L}$ von RIEMANNschen Flächen durch Streckenkomplexe darstellen, wenn man den Begriff des Streckenkomplexes allgemeiner faßt, als es bislang in der Regel geschehen ist[1]). Eine Fläche der Klasse $\mathfrak{L}$ kann nun weit allgemeinere Randstellen als logarithmische Windungspunkte tragen; als Beispiel mögen hier nur die mittelbaren Randstellen erwähnt werden, die bisher noch gar nicht syste-

*) Die Arbeit wurde von der Deutschen Forschungsgemeinschaft gefördert; ich sage dafür meinen Dank. Die Arbeit ist ein Teil der Habilitationsschrift des Verfassers, die der Naturwissenschaftlichen Fakultät der Justus-Liebig-Hochschule in Gießen vorgelegen hat.

[1]) HUCKEMANN [7]. Es gibt schon verschiedene Ansätze in dieser Richtung, die die vorliegende Arbeit gefördert haben. Vgl. hier auch Fußnote 2.

matisch von der geometrischen Seite her, ausgehend von der Flächenstruktur, behandelt und in ihrem Einfluß auf die Wertverteilung untersucht wurden. [Es liegt eine noch unveröffentlichte Untersuchung vor, die für die Lösungen linearer Differentialgleichungen die asymptotische Integration ausnutzt, um auf Flächenstruktur und Wertverteilung zu schließen[2]).] Unter den Flächen aus $\mathfrak{L}$ verdienen für unsere Zwecke besonders diejenigen hervorgehoben zu werden, deren (verallgemeinerter) Streckenkomplex aus endlich vielen periodischen Enden besteht, die einem Kernkomplex anhängen; entsprechend werden ja auch unter den Flächen aus $\mathfrak{R}$ im Zusammenhang mit Wertverteilungsfragen vorwiegend Flächen mit endlich vielen periodischen Enden betrachtet[3]).

Neben Randstellen beeinflußt auch Asymmetrie in den Enden die Wertverteilung der erzeugenden Funktion[4]); um den Einfluß der Randstellen möglichst rein herauszuschälen, wollen wir auf komplizierte Enden verzichten und unsere Betrachtungen beschränken auf Flächen mit logarithmischen Enden sowie ein- und mehrfachen Sinusenden.

Aus der Bauart einer Randstelle auf einer Fläche aus einer gewissen Unterklasse von $\mathfrak{L}$ können wir gewisse Zahlen gewinnen und der Randstelle als *Konvergenz- (Relativ-)* und *Verschmelzungsstärken* zuordnen; aus ihnen ergibt sich in eindeutiger Weise die *Wachstumsstärke* der Randstelle. Mit dem Begriff der Wachstumsstärke kann der bekannte *Randstellensatz*, nach dem eine Funktion wenigstens die Wachstumsordnung $n/2$ hat, wenn sie eine zerfällbare Fläche mit n unmittelbaren Randstellen erzeugt, wie folgt präzisiert werden: Die Wachstumsordnung ist *gleich* der halben Summe aus den Wachstumsstärken der Randstellen. Unmittelbare Randstellen zerfällbarer Flächen haben eine Wachstumsstärke ≥ 1, logarithmische Windungspunkte die genaue Wachstumsstärke 1; für die Stärke mittelbarer Randstellen gibt es dagegen keine positive untere Schranke, deren Stärke kann Null sein. Gewöhnliche Sinusenden (d.h. Sinusenden, deren Verzweigungspunkte nur über zwei Spurpunkten liegen) verhalten sich in ihrem Einfluß auf das Wachstum wie Randstellen der Stärke Null. Unsere Betrachtungen schließen auch gewisse Randstellen ein, die weder mittelbar noch unmittelbar sind; auch sie haben auf zerfällbaren Flächen eine Wachstumsstärke ≥ 1.

Auch der Defekt eines Wertes a hängt in einfacher Weise ab von den Konvergenz- und Verschmelzungsstärken der über a liegenden Randstellen, wobei allerdings, wie zu erwarten ist, auch die Stärken *aller* Randstellen mit hineinspielen. Zwischen dem Defekt $\delta(a)$ des Wertes a und dem Beitrag der über a liegenden Randstellen zur Wachstumsordnung besteht also eine gewisse Beziehung; sie äußert sich darin, daß positive Wachstumsstärke einer Randstelle über a einen positiven Wert von $\delta(a)$ nach sich zieht, und daß umgekehrt bei positivem $\delta(a)$ wenigstens *eine* Randstelle über a positive

[2]) MÖLLER-ULLRICH, Manuskript. Vgl. ULLRICH [25], KNESER-ULLRICH [8] Abschnitt C.

[3]) Die Bedeutung der periodischen Enden erkannte zuerst ULLRICH [23].

[4]) WITTICH [28].

Wachstumsstärke hat, wenn die Wachstumsordnung der Funktion nicht überhaupt Null ist. Damit folgt: sobald eine mittelbare Randstelle über a kräftig genug ist, einen Beitrag zur Wachstumsordnung zu liefern, ist auch $\delta(a)$ positiv: positiver Defekt ist sozusagen der Regelfall[5]). Es zeigt sich ferner, daß für endlich viele verschiedene Stellensorten die Defektverteilung im wesentlichen willkürlich vorschreibbar ist; dies führt über neuere von LE-VAN und PÖSCHL beim Umkehrproblem erzielte Ergebnisse hinaus[6]).

Wir verwenden zwei Methoden, die Wertverteilung zu ermitteln; eine der beiden geht auf AHLFORS[7]) zurück, die andere stützt sich auf den Modulsatz von TEICHMÜLLER[8]). Die erste kann man als gröber bezeichnen, denn sie liefert ungenauere Ergebnisse als die zweite, sie vermag aber auch noch bei Flächen, die sich vorerst der Methode der Modulabschätzung verschließen, Aussagen zu bringen. Damit wird auf eine Frage TEICHMÜLLERS[8]) nach der Stärke und Zugkraft beider Methoden eine Antwort gegeben.

§ 1. Netz und Streckenkomplex

Bekannt ist die Konstruktion des Streckenkomplexes für schlichtartige (einfach zusammenhängende) RIEMANNsche Flächen, deren Windungspunkte sich in endlich viele Grundpunkte projizieren, und der für sie eingeführte Streckenkomplexbegriff S[9]). Wir werden im folgenden einen allgemeineren Begriff benötigen[10]): unter einem Streckenkomplex S^* verstehen wir in dieser Arbeit eine Figur aus Knoten (Ringen und Kreuzen) sowie sie verbindenden Gliedern in der abgeschlossenen komplexen Ebene mit folgenden Eigenschaften:

1. *Die Anzahl der Knoten ist abzählbar unendlich.*
2. *Die Knoten und Glieder haben nur einen Häufungspunkt, und zwar den unendlich fernen Punkt U.*
3. *Jedes Glied verbindet einen Ring mit einem Kreuz.*
4. *Verschiedene Glieder haben außer den Knoten keine gemeinsamen Punkte.*
5. *Der Streckenkomplex ist zusammenhängend.*
6. *Von jedem Knoten gehen endlich viele Glieder aus, wenigstens aber zwei.*

Der Streckenkomplex zerlegt die Ebene in Teilgebiete; diejenigen unter ihnen, die keinen Knoten im Innern enthalten, heißen *Elementargebiete*; die

[5]) Eine Frage ULLRICHS [22], ob mittelbare Randstellen positiven Defekt bewirken können, wurde schon positiv beantwortet von MÖLLER (in einer unveröffentlichten Arbeit, vgl. ULLRICH [25]) und von TEICHMÜLLER [20].

[6]) Das Umkehrproblem der Wertverteilung gilt der Entscheidung folgender Frage: Gibt es gebrochene Funktionen, die eine vorgegebene Verteilung der Defekte und Verzweigungsindizes besitzen (ULLRICH [23])? Nach anderen partiellen Lösungen löste LE-VAN [10] das Umkehrproblem für den Fall, daß nur endlich viele Defekte und Verzweigungsindizes von Null verschieden und diese rationalzahlig sind. PÖSCHL [16] zeigte, daß auch gewisse algebraische Irrationalitäten als Defekte und Verzweigungsindizes auftreten können.

[7]) AHLFORS [1].

[8]) TEICHMÜLLER [19].

[9]) NEVANLINNA [14], [15] Kap. XI, § 2; ELFVING [2].

[10]) Vgl. [7]; die dortigen Ergebnisse werden in der hier benötigten Form in diesem Paragraphen zusammengestellt.

Glieder am Rand eines Elementargebietes heißen seine *Seiten*, die Knoten am Rand seine *Eckknoten*. Zwei Elementargebiete, die eine gemeinsame Seite haben, heißen *benachbart*, ebenso zwei Glieder, die beim Umlauf um einen Knoten unmittelbar aufeinander folgen. Wir fordern für S^* ferner

[6a. *Zwei Knoten sind nicht durch benachbarte Glieder verbunden.*]

7. *Jedem Elementargebiet ist eine reelle Zahl a als Koordinate zugeordnet*[11]).

8. *Zwei benachbarten Elementargebieten sind verschiedene Zahlen zugeordnet.*

9. *Wird ein Ring R einmal in positivem Sinn umlaufen, so folgen die den Elementargebieten mit dem Eckknoten R zugeordneten Zahlen in solcher Weise aufeinander, daß die reelle Achse genau einmal in positiver Richtung durchlaufen wird; das gleiche gilt beim Umlauf um ein Kreuz im negativen Sinn*[12]).

Jedem Ring wird eine obere, jedem Kreuz eine untere Halbebene zugeordnet. Sind zwei Knoten durch ein Glied verbunden, das zwei Elementargebiete der Koordinaten a_i, a_k trennt, so werden die zugeordnete obere und untere Halbebene über demjenigen von a_i, a_k begrenzten Intervall der reellen Achse [das den Punkt ∞ enthalten kann] verheftet, dessen Durchlaufung nach 9 dem Überschreiten des Gliedes entspricht. Dem Streckenkomplex S^* wird so eine nicht-kompakte RIEMANNsche Fläche W zugeordnet, die einfach zusammenhängend ist. Jedem $2n$-seitigen ($1 \leq n < \infty$) Elementargebiet der Koordinate a ist ein wohlbestimmter n-blättriger[13]) Windungspunkt über a, jedem unendlich-vielseitigen Elementargebiet der Koordinate a ist eine Randstelle von W über a zugeordnet. W hat Windungspunkte und Randstellen nur über der reellen Achse; umgekehrt läßt sich jede einfach zusammenhängende RIEMANNsche Fläche, die nur über der reellen Achse Windungspunkte und Randstellen trägt, und zwar in jedem Blatt in endlicher Anzahl, die ferner von keinem Teilbogen der reellen Achse berandet wird, durch einen Streckenkomplex S^* mit den Eigenschaften 1 bis 9 darstellen.

In jedem unendlich-vielseitigen Elementargebiet wird vom Punkt U aus, in jedem endlich-vielseitigen Elementargebiet von einem inneren Punkt P des Elementargebietes aus ein — bis eventuell auf die Endpunkte — im Elementargebiet verlaufender doppelpunktfreier JORDAN-Bogen so nach jedem Glied am Rand des Elementargebietes gezogen, daß die beiden auf beiden Seiten eines Gliedes endenden Bögen es im selben inneren Punkt treffen — beide zusammen nennen wir eine *Netzseite* — und daß verschiedene Bögen höchstens den Punkt U bzw. die Punkte P und den Schnittpunkt mit einem Gliede gemeinsam haben. Die Bögen selbst heißen *Netzkurven* oder *Netzlinien*.

[11]) In der in der Funktionentheorie üblichen Weise fassen wir ∞ als reelle Zahl auf.

[12]) Man kann natürlich hier und in 7. eine andere geschlossene doppelpunktfreie JORDAN-Kurve als die reelle Achse zugrunde legen; wir werden unsere Betrachtungen auf RIEMANNsche Flächen beschränken, die nur über der reellen Achse verzweigt sind. Die Wahl der reellen Achse an Stelle einer anderen geschlossenen doppelpunktfreien stückweise analytischen Kurve stellt keine wesentliche Einschränkung dar.

[13]) Ein einblättriger Windungspunkt ist ein schlichter Punkt; die Bedingung 6a hat $n \geq 2$ zur Folge, so daß mit 6a also jedem Elementargebiet ein eigentlicher (wenigstens zweiblättriger) Windungspunkt zugeordnet ist.

Jedes (eventuell erst im Unendlichen) geschlossene Polygon aus Netzlinien, das genau einen Knoten im Innern enthält, heißt *Elementarpolygon*. Die ganze von den Elementarpolygonen gebildete Figur F ist ein topologisches Bild von W derart, daß jedes Elementarpolygon topologisches Bild einer (oberen oder unteren) Halbebene, jede Netzseite topologisches Bild eines Stückes der reellen Achse ist, und daß der Punkt P in einem $2n$-seitigen ($n < \infty$) Elementargebiet der Koordinate a topologisches Bild eines n-blättrigen[13]) algebraischen Windungspunktes von W über a ist. F heißt ein *Netz* von W.

Die topologische Abbildung, die W in das Netz überführt, sei überdies so beschaffen, daß sie die Flächenpunkte über $w = i$ bzw. $w = -i$ in die (punktförmig idealisierten) Ringe bzw. Kreuze überführt und daß der Schnittpunkt einer Netzseite mit einem Glied $g(x_1, x_2)$, das zwei Elementargebiete der Koordinaten x_1, x_2 trennt, Bild des Punktes

$$(1) \qquad P(x_1, x_2) = P^{\pm}(x_1, x_2) = \frac{x_1 x_2 - 1 \pm \sqrt{x_1^2 x_2^2 + x_1^2 + x_2^2 + 1}}{x_1 + x_2}$$

über der reellen Achse ist; das Vorzeichen der Wurzel ist so zu bestimmen, daß $P(x_1, x_2)$ in dasjenige der beiden durch x_1, x_2 bestimmten Intervalle der reellen Achse fällt, über dem das Urbild der Netzseite liegt. An den Unbestimmtheitsstellen ist $P(x_1, x_2)$ stetig zu erklären[14]). Das Glied $g(x_1, x_2)$ sei Bild des Orthogonalkreisbogenstückes zur reellen Achse, das die Endpunkte $i, -i$ hat und $P(x_1, x_2)$ enthält.

Die in der Theorie des Streckenkomplexes S [s. Fußnote 9] übliche Terminologie (wie logarithmisches Elementargebiet, algebraisches Elementargebiet, logarithmisches Ende, periodisches Ende, Kern des Komplexes usw.) übernehmen wir wörtlich für S^*; sie bezieht sich ja nur auf den Streckenkomplex qua Streckenkomplex und nicht auf ihn als Darstellung einer RIEMANNschen Fläche.

Unter einem (einfachen) *Sinusende* verstehen wir das von einer Leiter aus vierseitigen Elementargebieten gebildete periodische Ende (Fig. 1a); die beiden Knotenfolgen am Rand heißen die *Holme* des Sinusendes; ein zwei Knoten verschiedener Holme verbindendes Glied heißt *Sprosse* des Sinusendes. Unter einem *n-fachen Sinusende* ($n \geq 1$) verstehen wir n so nebeneinander liegende Viereckleitern, daß zwei benachbarte (wenigstens) einen Holm gemeinsam haben[15]) (Fig. 1b für $n = 3$). Die Folge von Elementargebieten in einer der Viereckleitern heißt ein *Strang* des Sinusendes. Ein n-faches Sinusende hat also n Stränge. Logarithmische Elementargebiete und Stränge werden mit dem gemeinsamen Namen *Randgebiet* belegt, letztere werden auch

[14]) Bei Verschiebung eines Endpunktes des Intervalles bewegt sich $P(x_1, x_2)$ eigentlich monoton in derselben Richtung; dies ist hier die wesentliche Eigenschaft der Funktion $P(x_1, x_2)$.

[15]) Man beachte, daß zwei nebeneinanderliegende Viereckleitern auch beide Holme gemeinsam haben können; vgl. dazu ULLRICH [24], Fig. 15. Beispiele von Funktionen, die Flächen mit mehrfachen Sinusenden erzeugen, gibt PRATT [17].

Sinusgebiet genannt. Die beiden Randgebiete, die durch ein logarithmisches Ende oder einen Sinusholm getrennt werden, heißen die *Nachbargebiete* dieser Knotenfolge, beide Randgebiete selbst heißen *benachbart*. Wegen der zwischen der Fläche W und dem Netz bestehenden eineindeutigen Abbildung und weil die Glieder des Streckenkomplexes eindeutig auf Kurven der Fläche W bezogen sind, dürfen und werden wir die Bezeichnung Sinusgebiet auf die Punktmenge von W übertragen, die den Punkten des Sinusgebietes im Streckenkomplex entspricht.

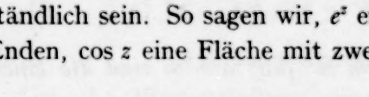


Fig. 1 a

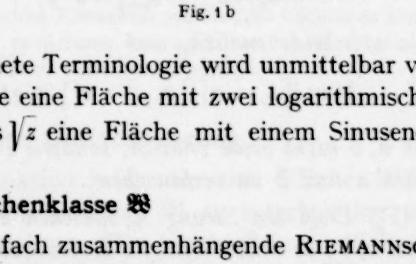


Fig. 1 b

Die in ausgearteten Fällen verwendete Terminologie wird unmittelbar verständlich sein. So sagen wir, e^z erzeuge eine Fläche mit zwei logarithmischen Enden, $\cos z$ eine Fläche mit zwei, $\cos\sqrt{z}$ eine Fläche mit einem Sinusende.

§ 2. Die Flächenklasse $\mathfrak{W}$

Wir betrachten nicht-kompakte einfach zusammenhängende RIEMANNsche Flächen, die Windungspunkte und Randstellen nur über der reellen Achse tragen; ihr Streckenkomplex S^* besitze einen endlichen Kern mit einer endlichen Anzahl periodischer Enden folgender Arten: logarithmische Enden, einfache Sinusenden, mehrfache Sinusenden. Es komme wenigstens ein Sinusende wirklich vor — für eine einfach zusammenhängende Fläche mit endlichem Kern sowie endlich vielen periodischen, und zwar nur logarithmischen, Enden im Streckenkomplex (die ja nur über endlich vielen Punkten verzweigt ist) wurde die Wertverteilung der erzeugenden Funktion bereits bestimmt[16]). Um die Bezeichnungsweise zu vereinfachen, soll jeder Strang eines mehrfachen Sinusendes auch einfach als *Sinusende* bezeichnet werden.

Die algebraischen Elementargebiete eines Sinusendes E zerfallen in zwei Klassen nach der Art der Aufeinanderfolge von Ring und Kreuz an der dem Kern zugewandten Seite des Elementargebiets. Die Elementargebiete der einen Klasse trennen die der anderen. Jede Sprosse in E entspricht als Glied des Streckenkomplexes einem bestimmten offenen Intervall i_ν ($\nu = 1, 2, \ldots$, gezählt vom Kern an) der reellen Achse. Wir nennen die Vereinigung $\bigcup_{\nu=1}^{\infty} i_\nu = h(E)$ die *Hülle*, den Durchschnitt $\bigcap_{\nu=1}^{\infty} i_\nu = k(E)$ den *Kern*[17]), die Menge $\bigcap_{n=1}^{\infty} \bigcup_{\nu=n}^{\infty} i_\nu = kh(E)$ den *Hüllenkern* des Sinusendes E und fordern:

(F_1) *Kein Sinusende habe als Kern die leere Menge.*

[16]) NEVANLINNA [12], [13], AHLFORS [1], WAGNER [27].

[17]) Die Begriffe „Kern des Streckenkomplexes" und „Kern des Sinusendes" sind auseinanderzuhalten. Der Zusammenhang läßt stets erkennen, welcher Begriff gemeint ist.

Für jedes Sinusende E werde ein Punkt des Kernes $k(E)$ ausgezeichnet, und zwar ein innerer Punkt, wenn $k(E)$ überhaupt innere Punkte hat; die Koordinate des ausgezeichneten Punktes heiße die *Koordinate des Sinusendes* bzw. des Stranges, Kernes. Es seien a, b die (notwendig verschiedenen) Koordinaten zweier benachbarter Stränge A, B in einem mehrfachen Sinusende. Die Koordinaten der Elementargebiete im Strang A bzw. B seien a_ν bzw. b_ν ($\nu = 1, 2, \ldots$, gezählt vom Kern des Streckenkomplexes an). Als eine gewisse Regelmäßigkeitsvoraussetzung über die Koordinaten a_ν, b_ν fordern wir:

(F_2) *Mit einer entweder konstanten oder unbeschränkten monotonen Funktion $g(\nu)$ und mit einer Konstanten K, die wir der Einfachheit halber positiv annehmen wollen, gelte für $\nu = 1, 2, \ldots$*

$$(F_2') \quad \operatorname{Max}\left\{\left|\frac{a_\nu - a}{a_\nu - b}\right|, \left|\frac{a_{\nu+1} - a}{a_{\nu+1} - b}\right|\right\} < g(\nu) + K < \operatorname{Min}\left\{\left|\frac{b_\nu - a}{b_\nu - b}\right|, \left|\frac{b_{\nu+1} - a}{b_{\nu+1} - b}\right|\right\}$$

falls a, b beide endlich, und

$$(F_2'') \quad \operatorname{Max}\{|a_\nu - a|, |a_{\nu+1} - a|\} < g(\nu) + K < \operatorname{Min}\{|b_\nu - a|, |b_{\nu+1} - a|\},$$

falls a, b nicht beide endlich, sondern etwa $b = \infty$ [falls $a = \infty$ sind die Buchstaben a und b zu vertauschen].

(F_3) *Liegt ein Strang A zwischen zwei Randgebieten verschiedener Koordinaten b, b', so sollen sich die Koordinaten der zu A gehörigen Windungspunkte weder gegen b noch gegen b' häufen.*

Es sei A ein Strang, B sei dasjenige Nachbargebiet von A, das — bei Blickrichtung zum Randpunkt ∞ von A — von A durch dessen linken Holm getrennt wird. $A' \neq A$ sei (wenn vorhanden) Nachbargebiet von B. Wir fordern dann:

(F_4) *Ist $h(A)$ die punktierte reelle Achse, so sei die Koordinate von A gleich der von A'.*

Die Forderung (F_1) und die Wahl der Koordinate eines Stranges haben zur Folge, daß in einem mehrfachen Sinusende die Hülle $h(A)$ eines Stranges A die Koordinate b eines benachbarten Stranges B nicht enthält; ferner sind genau dann nur $h(A)$ und b nicht durch ein Intervall der reellen Achse getrennt, wenn $h(A)$ die im Punkt b punktierte reelle Achse ist. Genau dann stimmen auch die Koordinaten der beiden Nachbargebiete des Stranges A überein. Die Forderungen (F_1) bis (F_4) sind natürlich in einer Untersuchung, die sich mit Randstellen beschäftigt, denn sie beseitigen zusätzliche Schwierigkeiten, die eine Kompliziertheit und Asymmetrie in der Struktur der periodischen Enden hervorrufen würde[18]).

Wir fassen die einfach zusammenhängenden Flächen mit endlichem Kern und endlich vielen logarithmischen Enden und Sinusenden im Streckenkomplex, die (F_1) bis (F_4) erfüllen, zur *Klasse* $\mathfrak{W}$ zusammen.

[18]) Man könnte (F_1), (F_2) dadurch abschwächen, daß man in den Definitionen von Hülle, Kern, Hüllenkern und in (F_2) ν an Stelle von 1 an von $N > 1$ an laufen ließe, ohne an den folgenden Überlegungen Wesentliches zu ändern.

Die Flächen $W \in \mathfrak{W}$ stehen in engem Zusammenhang mit den Flächen, die durch den Streckenkomplex S darstellbar sind[19]): sie gehen aus solchen dadurch hervor, daß die einzelnen algebraischen Windungspunkte auf der Zerschneidungskurve [hier: der reellen Achse] verschoben werden. Dies zeige folgende Skizze: Wir heben aus dem Streckenkomplex S^* einer Fläche $W \in \mathfrak{W}$ einen endlichen Teil T heraus, der den Kern des Komplexes ganz im Innern enthält. Den zwei Nachbargebieten eines Sinusendes E im übrigbleibenden Teil — seien es logarithmische Elementargebiete oder Sinusgebiete — ist je eine reelle Zahl, die Koordinate zugeordnet. Die Koordinaten der Windungspunkte von E fallen in dasjenige durch die Koordinaten der Nachbargebiete bestimmte offene Intervall der reellen Achse (das die punktierte reelle Achse ist, wenn beide Koordinaten übereinstimmen), das die Koordinate von E enthält; diese selbst trennt die Koordinaten der beiden Klassen von Windungspunkten in E. Verschiebt man über der Zerschneidungskurve die algebraischen Windungspunkte, die algebraischen Elementargebieten außerhalb T zugeordnet sind, so kann man erreichen, daß diejenigen Windungspunkte, die den außerhalb T liegenden algebraischen Elementargebieten irgendeines Sinusendes zugeordnet sind, über zwei Koordinaten zu liegen kommen. Die Zahl der Elementargebiete innerhalb T ist endlich, ebenso die Zahl der logarithmischen Elementargebiete (die höchstens gleich der Zahl der periodischen Enden ist); somit projizieren sich nach solcher Verschiebung sämtliche Windungspunkte der Fläche in endlich viele Grundpunkte.

§ 3. Zerschneidung und Uniformisierung

Um die Wertverteilung einer Funktion $w(z)$, die eine Fläche $W \in \mathfrak{W}$ erzeugt, zu bestimmen, erscheint es zweckmäßig, W zu zerschneiden und die einzelnen Teile getrennt zu uniformisieren. Wir beschreiben das Schnittsystem der Übersichtlichkeit halber im schlichten Netz-Streckenkomplex-Bild.

In jedem periodischen Ende gibt es *erste Knoten*, d.h. Knoten, die unmittelbar durch *ein* Glied mit einem Knoten des Streckenkomplex-Kerns verbunden sind (in den Ausartungsfällen der Streckenkomplexe von e^z, $\cos z$ usw. wird der Kern des Komplexes definitorisch so festgelegt, daß der Kern wenigstens einen Knoten enthält); letztere seien selbst als *nullte* Knoten des periodischen Endes bezeichnet. Logarithmische Enden haben *einen*, n-fache Sinusenden haben $n+1$ erste Knoten. Wir schneiden jedes periodische Ende nun längs derjenigen Netzlinien ab, die die ersten Knoten vom Kern des Komplexes trennen. Ferner werde in jedem Strang eines Sinusendes ein Schnitt auf allen Netzseiten gezogen, die „parallel zu den Holmen" von einem Elementargebiet zum andern im Strang laufen. Schließlich werden Schnitte gezogen längs der Glieder am Rand des Kerns, die nullte Knoten periodischer Enden verbinden; sie seien von dem nullten Knoten eines Endes aus, in den ein Schnittglied mündet, auf dem ersten benachbarten Glied des periodischen Endes noch so weit verlängert, daß sie die vorher gezogenen Schnitte gerade treffen. Fig. 2 zeigt Streckenkomplex und Schnittkurven (stark ausgezogen) für eine Fläche mit zwei logarithmischen Enden, zwei einfachen und einem vierfachen Sinusende.

Die endlich vielen relativ zu W abgeschlossenen Teilflächenstücke f, in die W durch die Zerschneidung zerfällt, gehören einer der folgenden drei Arten F, G, H an: Es sei

[19]) Hier können sogar (F_2), (F_3), (F_4) fallengelassen werden.

$f \in F$, wenn das Bild von f im Netz ganz im endlichen liegt, f somit ein kompaktes endlich viele Blätter enthaltendes Flächenstück ist;

$f \in G$, wenn f ein endlich-vielblättriges nicht-kompaktes Flächenstück ist, dessen Bild im Netz in ein logarithmisches Elementargebiet hinein zum Punkt U reicht;

$f \in H$, wenn f ein unendlich-vielblättriges Flächenstück ist.

f^0 bezeichne die Menge der inneren Punkte von f.

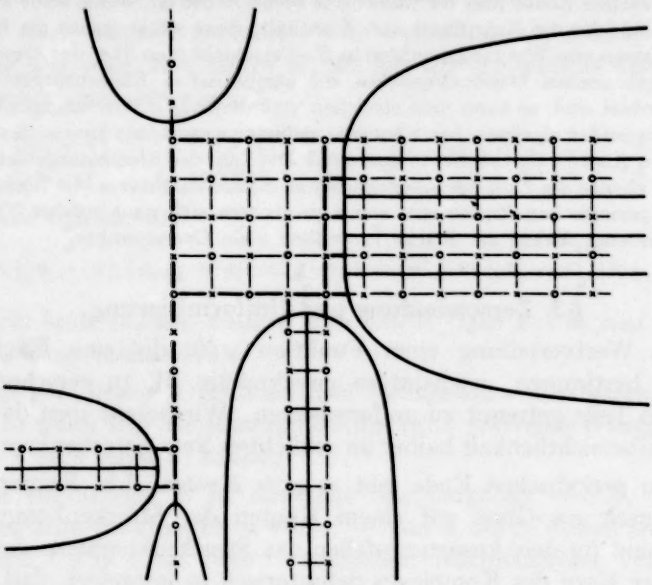


Fig. 2

Die Flächenstücke $f \in F \cup G \cup H$ können in beliebiger endlicher Anzahl auftreten. Die Flächenstücke der Klasse F werden uns im folgenden nicht beschäftigen, eine untergeordnete Rolle spielen auch die Flächenstücke $f \in G$. Wichtig sind die zuletzt genannten Flächenstücke $f \in H$, denn ihre durch die Randstellen bestimmte Struktur legt das Wertverteilungsverhalten von $w(z)$ fest.

Zwei Flächenstücke f, f', die beide in $G \cup H$ enthalten sind, heißen *benachbart*, wenn ihre Bilder im Netz längs eines Stückes einer nach Unendlich laufenden Netzseite aneinander stoßen. Die Gesamtheit der gemeinsamen Randpunkte heißt *rechter* bzw. *linker Rand* für dasjenige Flächenstück, das zur Linken bzw. zur Rechten liegt, wenn der gemeinsame Rand so durchlaufen wird, daß die Bildkurve im Netz in der zum Punkt U führenden Richtung orientiert ist.

Das Bild eines Flächenstückes $f \in H$ im Netz enthält eine der folgenden drei charakteristischen Knotenfolgen des Streckenkomplexes:

a) *Die Knotenfolge eines logarithmischen Endes.* Die benachbarten Randgebiete sind logarithmische Elementargebiete;

b) *die Knotenfolge am Rand eines (ein- oder mehrfachen) Sinusendes.* Die beiden Nachbargebiete sind ein logarithmisches Elementargebiet und ein Sinusgebiet;

c) *die Knotenfolge im Innern eines mehrfachen Sinusendes.* Die Nachbargebiete sind Sinusgebiete.

Wir fassen die Flächenstücke, die diesen drei Typen von Knotenfolgen entsprechen, in die bezüglichen Klassen H_a, H_b, H_c, so daß $H = H_a \cup H_b \cup H_c$. Den Nachbargebieten einer jeden solchen unendlichen Knotenfolge — seien es nun logarithmische Elementargebiete oder Sinusgebiete — sind zwei verschiedene Zahlen w' und w'' als Koordinaten zugeordnet; wie man sieht, liegen dann weder über w' noch über w'' innere Punkte des der Knotenfolge zugehörigen Flächenstückes; also ist $\log \frac{w - w'}{w - w''}$ eine im Innern des Flächenstückes eindeutige Funktion — die selbstverständliche kleine Modifikation, die hier oder an ähnlicher Stelle notwendig wird, wenn w' oder w'' unendlich sind, will ich übergehen —. Wir sagen, $f \in H$ habe ein Sinusgebiet oder Sinusende bzw. ein logarithmisches Gebiet (links oder im linken Teil, rechts oder im rechten Teil), wenn die zugehörige Knotenfolge als (linkes, rechtes) Nachbargebiet ein Sinusgebiet bzw. ein logarithmisches Elementargebiet besitzt. Zur Schreibweise verabreden wir: g, h, h_a, h_b, h_c bedeuten immer Flächenstücke aus G, H, H_a, H_b, H_c beziehentlich; Flächenstücke f können zu einer beliebigen Klasse gehören; $f^0, g^0, h^0, h_a^0, h_b^0, h_c^0$, bezeichnen die Menge innerer Punkte in den entsprechenden Flächenstücken. Ein Flächenstück g bzw. h_a, h_b, h_c hat als Nachbarflächenstücke nur Flächenstücke h_a oder h_b bzw. g, g oder h_b oder h_c, h_b oder h_c. In jedem Flächenstück g ist eine Randstelle von W erreichbar; ist a die Koordinate dieser Randstelle, so ist wegen der Wahl der topologischen Abbildungen von W ins Netz kein innerer Punkt von g über $P^-(a, a) = -a^{-1}$ enthalten; es ist also $\log \frac{w - a}{w + a^{-1}}$ für $a \neq 0, \infty$, und $\log w$ für $a = 0$ oder $a = \infty$ in g^0 eindeutig.

Ein Flächenstück $f \in G \cup H_b$ enthält am Rand genau eine Randstelle von W, die *unmittelbar relativ zu* f ist; sie entspricht dem logarithmischen Elementargebiet, in das das Netzbild von f hineinreicht. $f \in H_b$ hat genau dann eine weitere Randstelle von W am Rand, die *mittelbar ist relativ zu* f, wenn der Hüllenkern des Sinusendes von $f = h_b$ aus nur einem Punkt besteht. Ebenso enthält ein Flächenstück h_c am Rand null, eine oder zwei (relativ mittelbare) Randstellen von W, je nachdem sich null, einer oder zwei Hüllenkerne der Nachbargebiete für die zugehörige Knotenfolge auf einen Punkt reduzieren[20]). Ein Flächenstück h_a enthält am Rand stets zwei verschiedene (relativ unmittelbare) Randstellen von W, die den logarithmischen Elementargebieten entsprechen, die Nachbargebiete der zugehörigen Knotenfolge sind. Da jedes

[20]) Jeder Weg, dessen Bild im Netz in einem Strang E gegen den Punkt U strebt, muß sich von der Koordinate eines inneren Punktes des Hüllenkernes $kh(E)$ immer wieder um ein positiv nach unten beschränktes Stück entfernen; andererseits hat der Weg auf der holmparallelen Netzlinie in E ein Bild auf W, dessen Koordinaten sich schließlich beliebig wenig von $kh(E)$ entfernen.

Sinusende und jedes logarithmische Elementargebiet Nachbargebiet für zwei Knotenfolgen ist, muß die Anzahl der relativen Randstellen in den Flächenstücken wenigstens doppelt so groß sein, wie die Anzahl der Randstellen von W — sie kann aber noch wesentlich größer sein, wenn nämlich mehreren Randgebieten dieselbe Randstelle von W zugeordnet ist, wenn, wie wir sagen, verschiedene Relativ-Randstellen zu einer einzigen Randstelle der Fläche verschmolzen sind[21]). Fig. 3 gibt ein typisches Beispiel: sie zeigt den Streckenkomplex einer Fläche, in der zwei mittelbare (mit Bildern in den Randgebieten R_2, R_3) und zwei unmittelbare Relativ-Randstellen (mit Bildern in den Randgebieten R_1, R_4) zu einer einzigen (weder mittelbaren noch unmittelbaren) Randstelle der Fläche verschmolzen sind.

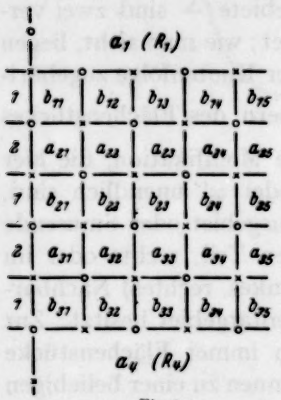


Fig. 3.
Die Koordinaten sind in die Elementargebiete hineingeschrieben. $a_1 = a_4 = \infty$; $b_{1\nu} = b_{2\nu} = b_{3\nu} = (-1)^\nu \cdot \left[\frac{\nu+2}{2}\right]$; $a_{2\nu} = a_{3\nu} = (-1)^\nu \cdot \left[\frac{\nu+4}{2}\right]$

§ 4. Die Ausschöpfung der Fläche

Wir nehmen an, daß bei der im letzten Paragraphen beschriebenen Zerschneidung m Flächenstücke $f_\mu \in G \cup H$ ausfallen ($\mu = 1, 2, \ldots, m$) — diese werden unserer Verabredung gemäß auch mit g, h, h_a, h_b, h_c respektive bezeichnet. Es hatte sich ergeben, daß es für jedes Flächenstück f_μ zwei Zahlen w'_μ, w''_μ derart gibt, daß $\zeta_\mu(w) = \log \frac{w - w'_\mu}{w - w''_\mu}$ — mit trivialer Modifikation, falls w'_μ oder w''_μ unendlich — in f^0_μ eine eindeutige Funktion darstellt. Für $f_\mu \in H$ sei w'_μ die Koordinate des linken, w''_μ die Koordinate des rechten Nachbargebietes der f_μ zugehörigen Knotenfolge; für $f_\mu \in G$ sei w''_μ die Koordinate der in f_μ erreichbaren Randstelle von W und $w'_\mu = -w''^{-1}_\mu$. $\zeta_\mu(w)$ bildet ein Flächenstück g^0 ab in einen Halbstreifen der Breite $n\pi$, wenn n die Anzahl der Knotenpunkte am Rand des Bildes von g im Netz ist; das ζ_μ-Bild eines Flächenstückes h^0 liegt in einer Halbebene $\Im(\zeta_\mu) > \gamma_\mu$, γ_μ reell. Als Bild von h^0_a entsteht eine Halbebene; das Bild von h^0_b ist eine Halbebene, die auf Parallelen zur reellen Achse im Abstand π Einschnitte, *Gräten*, trägt, sie rühren davon her, daß die im Sinusgebiet verlaufende Berandung von h_b mit einem Punkt über der Koordinate des Sinusendes ein ganzes Intervall der reellen Achse mit der Koordinate im Innern enthält [vgl. (F_1)]. Die Halbebene ist entweder nach rechts oder nach links aufgeschnitten. Ganz ähnliche Verhältnisse liegen vor bei dem ζ_μ-Bild von h^0_c; der einzige Unterschied ist, daß die Bildhalbebene sowohl nach rechts wie nach links auf Parallelen zur reellen Achse im Abstand π aufgeschlitzt ist, also beidseitig Gräten trägt.

Durch eine Transformation $\zeta^{(\mu)}(w) = \zeta_\mu(w) + i\gamma^{(\mu)}$, $\gamma^{(\mu)}$ reell, erreichen wir, daß die $\zeta^{(\mu)}$-Bilder $h^{0\zeta}$ von h^0 mit der oberen Halbebene $\Im(\zeta^{(\mu)}) > 0$ bzw. mit der

[21]) Genaueres darüber in § 7 und in [7]. Zur Klassifikation von Randstellen vgl. auch IVERSEN [9] und ULLRICH [24].

geschlitzten oberen Halbebene zusammenfallen; ebenso erreichen wir, daß das $\zeta^{(\mu)}$-Bild $g^{0\zeta}$ eines g^0, dessen Bild im Netz n Knotenpunkte am Rand enthält, in $0<\Im(\zeta^{(\mu)})<n\pi i$, $\Re(\zeta^{(\mu)})>\xi^{(\mu)}$ mit geeignetem reellen $\xi^{(\mu)}$ fällt (Fig. 4). Die Begrenzungen der Bildgebiete $h^{0\zeta}$, $g^{0\zeta}$ sind in gewisser Weise miteinander zu identifizieren, wie bestimmt durch das Aneinanderstoßen der Flächenstücke h, g auf der Fläche W. Wir beachten, daß die Anzahl der Flächenstücke $h\in H$ größer oder gleich 2 ist, da W wenigstens ein Sinusende hat [aus dem bei der Zerschneidung zwei Flächenstücke $h\in H$ oder Teile von solchen hervorgehen]. Es werde ein Flächenstück $f_1\in G\cup H$ als *erstes* ausgezeichnet; das über den rechten Rand an f_μ anstoßende rechte Nachbarflächenstück von f_μ bezeichnen wir mit $f_{\mu+1}$ ($\mu=1,2,\ldots,m$; genommen modulo m). Das $\zeta^{(\mu)}$-Bild Q_μ von f_μ^0 werde in die Ebene E_μ mit der laufenden Koordinate $\zeta_\mu=\xi_\mu+i\eta_\mu$ gelegt. Wir bezeichnen $E_{\mu+1}$ auch als *rechte Nachbarebene* von E_μ. Die durch ein linkes (rechtes) Sinusgebiet eines Flächenstückes $f_\mu\in H_b\cup H_c$ hervorgerufenen Gräten im $\zeta^{(\mu)}$-Bild heißen *linke (rechte) Gräten*, das $\zeta^{(\mu)}$-Bild des linken (rechten) Randes von $f_\mu\in G\cup H$ heiße *linker (rechter) Rand* von Q_μ.

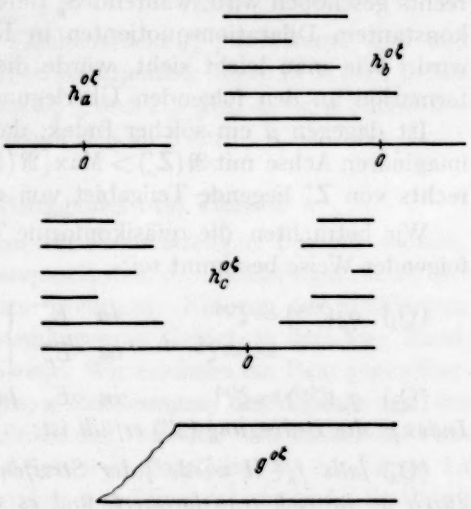


Fig. 4

In den uniformisierenden Transzendenten $\zeta^{(\mu-1)}(w)$, $\zeta^{(\mu)}(w)$ ist

$$w'_\mu=w''_{\mu-1}, \quad \text{wenn } f_\mu\in H, \tag{2}$$

$$w''_\mu=-\frac{1}{w'_\mu}=w''_{\mu-1}, \quad \text{wenn } f_\mu\in G; \tag{2'}$$

es ist überdies jedenfalls

$$w''_\mu=w'_{\mu-1}, \tag{2''}$$

wenn $f_\mu, f_{\mu-1}$ Teile eines Sinusgebietes gemeinsam haben, dessen Hülle die punktierte reelle Achse ist.

Es sei zunächst $f_\mu\in H$ und es sei (2'') nicht erfüllt. Dann gibt es einen Streifen $K_\mu^-<\Re(\zeta_\mu)<K_\mu^+$ in E_μ, der nur endlich viele Gräten trifft. [Nach (F_4) hat kein Sinusende von f_μ als Hülle die punktierte reelle Achse, nach (F_3) häufen sich die Realteile rechter bzw. linker Grätenendpunkte nicht gegen $-\infty$ bzw. $+\infty$, womit nach (F_2) die Behauptung folgt.] Ist dazu $f_{\mu-1}\in H$, so folgen die Punkte $\zeta^{(\mu)}(w'_{\mu-1})$ auf einer Parallelen zur imaginären Achse im Abstand 2π aufeinander, sie haben — von etwa vorhandenen — linken Gräten positiven Abstand. $\mathfrak{Z}_\mu^-$, $\mathfrak{Z}_\mu^+$ seien zwei Parallelen zur imaginären Achse innerhalb des Streifens mit $\Re(\mathfrak{Z}_\mu^-)<\Re\left(\zeta^{(\mu)}(w'_{\mu-1})\right)$, $\Re(\mathfrak{Z}_\mu^-)<\Re(\mathfrak{Z}_\mu^+)$. E_μ^-

sei das links von $\mathfrak{Z}_\mu^-$, E_μ^+ das rechts von $\mathfrak{Z}_\mu^+$, S_μ das zwischen $\mathfrak{Z}_\mu^-$ und $\mathfrak{Z}_\mu^+$ liegende Teilgebiet der oberen Halbebene. Wir dürfen für das folgende annehmen, daß auch das durch $-\zeta^{(\mu-1)}$ vermittelte Bild der $\mathfrak{Z}_\mu^-$ auf W entsprechenden Kurve ganz links von $\mathfrak{Z}_\mu^+$ liegt [vgl. $(\mathfrak{Q}_1)$, $(\mathfrak{Q}_3)$ unten]; ist dies nämlich nicht sowieso der Fall, so würde man E_μ so quasikonform auf sich abbilden, daß E_μ^- ungeändert bleibt, daß E_μ^+ durch eine Translation genügend weit nach rechts geschoben wird, während S_μ durch eine quasikonforme Abbildung mit konstantem Dilatationsquotienten in Richtung der reellen Achse gestreckt wird. Wie man leicht sieht, würde diese zusätzliche quasikonforme Transformation an den folgenden Überlegungen nichts Wesentliches ändern.

Ist dagegen μ ein solcher Index, daß $f_\mu \in G$, so sei $\mathfrak{Z}_\mu^+$ eine Parallele zur imaginären Achse mit $\Re(\mathfrak{Z}_\mu^+) > \mathrm{Max}\left[\Re\left(\zeta^{(\mu)}(w'_{\mu-1})\right), \Re\left(\zeta^{(\mu)}(w''_{\mu+1})\right)\right]$ und E_μ^+ das rechts von $\mathfrak{Z}_\mu^+$ liegende Teilgebiet von Q_μ.

Wir betrachten die quasikonforme Transformation $q_\mu(\zeta_\mu)$ von E_μ, die in folgender Weise bestimmt sei:

$(\mathfrak{Q}_1)$ $\left.\begin{aligned} q_\mu(\zeta^{(\mu)}) &= \zeta^{(\mu)} \quad \text{in} \quad E_\mu^+ \\ &= -\zeta^{(\mu-1)} \quad \text{in} \quad E_\mu^- \end{aligned}\right\}$ *falls* $f_\mu \in H$, $f_{\mu-1} \in H$;

$(\mathfrak{Q}_2)$ $q_\mu(\zeta^{(\mu)}) = \zeta^{(\mu)}$ *in* E_μ, *falls* $f_\mu \in H$, $f_{\mu-1} \in G$, *oder falls für den Index* μ *die Bedingung* (2″) *erfüllt ist;*

$(\mathfrak{Q}_3)$ *falls* $f_\mu \in H$ *werde jeder Streifen* $(\nu-1)\pi \leq \Im(\zeta_\mu) \leq \nu\pi$ $(\nu = 1, 2, \ldots)$ *durch* q_μ *in sich transformiert, und es sei außerdem* q_μ *schlicht und zweimal stetig differenzierbar in* E_μ *und habe die Periode* $2\pi i$;

$(\mathfrak{Q}_4)$ *falls* $f_\mu \in G$, *sei*

$$q_\mu(\zeta^{(\mu)}) = \frac{2}{n\pi}\left[\left(\eta_\mu - \frac{n\pi}{2}\right)\zeta^{(\mu-1)} + (n\pi - \eta_\mu)\zeta^{(\mu)}\right] \text{ in } E_\mu^+ \text{ für } \frac{n\pi}{2} \leq \eta_\mu \leq n\pi$$

$$= \frac{2}{n\pi}\left[-\left(\frac{n\pi}{2} - \eta_\mu\right)\zeta^{(\mu+1)} + \eta_\mu \cdot \zeta^{(\mu)}\right] \text{ in } E_\mu^+ \text{ für } 0 \leq \eta_\mu \leq \frac{n\pi}{2},$$

wobei $n\pi$ *der Abstand des rechten und linken Randes von* Q_μ *ist.*

Falls $f_\mu \in H$, ist der Dilatationsquotient der Abbildung q_μ nur in $|\Re(\zeta_\mu)| < K$, mit geeignetem $K > 0$, von Eins verschieden und dort beschränkt; falls $f_\mu \in G$, so ist der Dilatationsquotient in Q_μ beschränkt. Ferner erhält man

$$q_\mu(\zeta_\mu) - \zeta_\mu = O(1). \tag{3}$$

Die Transformation q_μ bewirkt weiter, daß zwei miteinander zu identifizierende Punkte auf dem rechten Rand von $q_{\mu-1}(Q_{\mu-1})$ und dem linken Rand von $q_\mu(Q_\mu)$, die genügend weit vom Nullpunkt entfernt sind, entweder zur imaginären Achse spiegelbildlich liegen (wenn $f_\mu \in H$) oder durch die Transformation $q_\mu - n\pi i$ zur Deckung gebracht werden (wenn $f_\mu \in G$, wobei n die Knotenanzahl am Rand des Bildes von f_μ im Netz ist).

Für hinreichend großes ϱ gibt es auf dem Kreisbogen $|\zeta_\mu| = \varrho$ für $f_\mu \in H$, bzw. auf der Geraden $\Re(\zeta_\mu) = \varrho$ für $f_\mu \in G$, ein wohlbestimmtes Stück B_ϱ^μ resp. $\tilde{B}_\varrho^\mu$, das den rechten und den linken Rand von $q_\mu(Q_\mu)$ resp. Q_μ verbindet und

bis auf die Endpunkte ganz innerhalb $q_\mu(Q_\mu)$ resp. Q_μ liegt[22]). Auf Grund der Bestimmung von $q_{\mu-1}$, q_μ ist für hinreichend großes ϱ der Schnittpunkt von $B_\varrho^{\mu-1}$ mit dem rechten Rand von $q_{\mu-1}(Q_{\mu-1})$ äquivalent zum Schnittpunkt von B_ϱ^μ mit dem linken Rand von $q_\mu(Q_\mu)$, d. h. die Urbilder beider Punkte auf der Fläche W sind miteinander zu identifizieren: sie fallen zusammen. Es ist für solche ϱ also das Bild $B_{\varrho,w}$ von $\bigcup_{\mu=1}^{m} B_\varrho^\mu$ auf W eine geschlossene Kurve. die überdies doppelpunktfrei ist, weil die Flächenstücke f_μ^0 punktfremd sind und höchstens je zwei benachbarte von ihnen Randpunkte des rechten oder linken Randes gemeinsam haben. Zwei Kurven $B_{\varrho',w}$ und $B_{\varrho'',w}$ $(\varrho_0<\varrho'<\varrho'')$ schließen also für genügend großes ϱ_0 ein Ringgebiet $G(\varrho', \varrho'')$ ein.

§ 5. Der Modul von Ringgebiet und Viereck

Ein Ringgebiet kann stets konform auf einen Kreisring bezogen werden; der positive Logarithmus des Radienquotienten des Kreisringes heißt der *Modul* des Ringgebietes, er ist konform invariant. Ein von der punktierten Ebene verschiedenes einfach zusammenhängendes Gebiet, in dem vier Randelemente ausgezeichnet sind, heißt *Viereck*. Wir zeichnen ein Paar gegenüberliegender Seiten, die *α-Seiten (α-Mengen, α-Randmengen)* des Vierecks aus; das andere Paar bezeichnen wir als die *β-Seiten* des Vierecks. Wir bilden das Viereck so in ein Rechteck ab, daß die β-Seiten in Parallelseiten der Länge 2π übergehen; die Länge M der transformierten α-Seiten ist eindeutig bestimmt und heißt *Modul des Vierecks* [bezüglich der ausgewählten α-Seiten; Vertauschung von α- und β-Seiten führt zum Modul $4\pi^2/M$]. Wird ein Kreisring vom Modul M durch einen Schnitt auf einem Radius in ein Viereck zerlegt, und werden die Schnittufer als α-Seiten gewählt, so hat das gewonnene Viereck ebenfalls den Modul M.

Es gelten folgende Lemmen[23]).

Lemma 1. *Wird ein Viereck (Ringgebiet) über einer β-Seite (einem Randkontinuum) erweitert oder über einer α-Seite verengt, so steigt der Modul*

Lemma 2. *Der Modul eines Viereckes steigt, wenn ein Teil der β-Seiten als α-Menge gezählt wird, d. h. wenn ein Eckpunkt des Vierecks so auf der Berandung verschoben wird, daß die α-Menge vermehrt wird.*

Lemma 3. a) *Liegen in einem Kreisring R des Moduls M zwei fremde Ringgebiete R', R'' der Moduln M', M'', von denen jedes die beiden Randkontinuen trennt, so gilt $M'+M''\leq M$; das Gleichheitszeichen steht genau dann, wenn R', R'' konzentrische Kreisringe sind, die zusammen den Kreisring R ausfüllen.*

[22]) Dies ist unmittelbar einzusehen falls $f_\mu \in G \cup H_a \cup H_b$. Das gleiche gilt für $f_\mu \in H_c$, falls $q_\mu(\zeta^{(\mu)}) = \zeta^{(\mu)}$ ist [jedenfalls also wenn ein Sinusende von f_μ als Hülle die punktierte reelle Achse hat], ebenso, falls es auch in $q_\mu(Q_\mu)$ einen Vertikalstreifen gibt, der nur endlich viele Gräten trifft. Andernfalls erkennt man für $f_\mu \in H_c$ die Richtigkeit der Aussage daraus, daß bei einer Bewegung auf dem Kreis $|\zeta_\mu| = \varrho$ vom Punkt $i\varrho$ aus zu einem Punkt mit dem Imaginärteil $\varrho - 1$ sich der Realteil in der Größenordnung $\sqrt{\varrho}$ ändert.

[23]) Sie finden sich in TEICHMÜLLER [19] oder sind mit den dortigen Überlegungen leicht zu gewinnen.

b) *Wird ein Kreisring des Moduls M durch beide Randkontinuen verbindende punktfremde Querschnitte in n Vierecke zerlegt, die bezüglich der Querschnitte als α-Seiten die Moduln M_ν $(\nu = 1, 2, \ldots, n)$ haben, so gilt* $\sum_{\nu=1}^{n} M_\nu^{-1} \leq M^{-1}$; *das Gleichheitszeichen steht genau dann, wenn alle Querschnitte auf Radien des Kreisringes liegen.*

Lemma 4 [24]). *Es werde ein Rechteck R des Moduls M in der z-Ebene $(z = x + iy)$, dessen β-Seiten Parallelseiten der Länge 2π sind, quasikonform auf ein Viereck V' des Moduls M' abgebildet; $Q(z)$ sei der Dilatationsquotient an der Stelle z. Dann gilt*

$$|M - M'| \leq \frac{1}{2\pi} \iint_R \{Q(z) - 1\}\, dx\, dy.$$

Lemma 5. *In der z-Ebene sei gegeben eine Kurvenschar C_ϱ $(\varrho > 0)$ von geschlossenen doppelpunktfreien Kurven mit der Eigenschaft, daß für $\varrho' < \varrho''$ die Kurve $C_{\varrho'}$ durch die Kurve $C_{\varrho''}$ vom Punkt ∞ getrennt wird und daß C_ϱ sich für $\varrho \to \infty$ auf den Punkt ∞ zusammenzieht. Der Modul des von $C_{\varrho'}$, $C_{\varrho''}$ eingeschlossenen Ringgebietes sei $M(\varrho', \varrho'')$. Wir setzen*

$$u_1(\varrho) = \operatorname*{Min}_{C_\varrho} |z|, \qquad u_2(\varrho) = \operatorname*{Max}_{C_\varrho} |z|,$$

und bezeichnen Größen, die im limes $\varrho \to \infty$ verschwinden, mit $\varepsilon(\varrho)$. Wenn dann eine Funktion $g(\varrho)$ und eine Zahl $K > 0$ existiert, so daß mit festem $\varrho_0 > 0$ für $\varrho_0 < \varrho' < \varrho''$ und $M(\varrho', \varrho'') \geq K$ stets die Beziehung

$$M(\varrho', \varrho'') - \int_{\varrho'}^{\varrho''} g(\varrho)\, \frac{d\varrho}{\varrho} = \varepsilon(\varrho')$$

gilt, so ist mit geeignetem $k > 0$

$$u_i(\varrho) = k e^{f(\varrho)} \left(1 + \varepsilon(\varrho)\right), \quad \text{wo} \quad f(\varrho) = \int_{\varrho_0}^{\varrho} g(\varrho)\, \frac{d\varrho}{\varrho}, \quad i = 1, 2.$$

In anderen Worten: *unter den Voraussetzungen des Lemma 5 sind die Kurven C_ϱ asymptotisch kreisnahe.* Bei der Untersuchung von Wertverteilungsproblemen ist es oft nützlich, solche Kurven zur Verfügung zu haben.

§ 6. Bestimmung der Wachstumsordnung

6.1. Wir nehmen in diesem ganzen Paragraphen an, daß $\varrho, \varrho', \ldots$ größer seien als eine hinreichend groß gewählte Zahl; $\varepsilon(x)$ bezeichne Größen, die im limes $x \to \infty$ verschwinden.

Der Modul des Ringgebietes $G(\varrho', \varrho'')$, das von den Kurven $B_{\varrho', w}$ und $B_{\varrho'', w}$ $(\varrho' < \varrho'')$ eingeschlossen wird, sei $M(\varrho', \varrho'')$. Durch die Schnittkurven σ auf der Fläche W, die — gleich wie ihre konformen und quasikonformen Bilder in den Ebenen E_μ — immer als α-Menge gewählt werden, wird $G(\varrho', \varrho'')$ in m

[24]) AF HÄLLSTRÖM [4] beweist das entsprechende Lemma für Ringgebiete.

verschiedene Vierecke zerlegt; der Modul desjenigen Vierecks, das durch $\zeta^{(\mu)}(w)$ in die Ebene E_μ abgebildet wird, sei $M_\mu(\varrho', \varrho'')$, der Modul des durch q_μ vermittelten quasikonformen Bildes $v_\mu(\varrho', \varrho'')$ sei $m_\mu(\varrho', \varrho'')$. Nach Lemma 3b) gilt

$$\sum_{\mu=1}^{m} \frac{1}{M_\mu(\varrho', \varrho'')} \leq \frac{1}{M(\varrho', \varrho'')}. \tag{4}$$

Die Länge von B_ϱ^μ sei $\pi\lambda_\mu(\varrho)\cdot\varrho$; $\lambda_\mu(\varrho)$ ist eine stückweise stetige Funktion. Falls ϱ_μ eine Stetigkeitsstelle von $\lambda_\mu(\varrho)$ ist [dann trifft jedenfalls B_ϱ^μ den Rand von $q_\mu(Q_\mu)$ unter von Null verschiedenem Winkel], erhält man mit Hilfe der Lemmata 1, 2 die Gleichung

$$m_\mu(\varrho_\mu, \varrho_\mu + d\varrho) = \frac{2}{\lambda_\mu(\varrho_\mu)} \frac{d\varrho}{\varrho_\mu}. \tag{5}$$

Mit Lemma 4 folgt weiter die Gleichung

$$M_\mu(\varrho_\mu, \varrho_\mu + d\varrho) = m_\mu(\varrho_\mu, \varrho_\mu + d\varrho) + \frac{O(1)}{\{\lambda_\mu(\varrho_\mu)\}^2} \cdot \frac{d\varrho}{\varrho_\mu^2}, \tag{6}$$

denn für $f_\mu \in H$ ist der Dilatationsquotient der Transformation q_μ beschränkt und nur in einem Vertikalstreifen endlicher Breite von Eins verschieden, während er für $f_\mu \in G$ beschränkt ist. Es sei $\lambda_\mu(\varrho)$ stetig für alle μ an einer Stelle ϱ_0. Dann gilt für das infinitesimale Ringgebiet $G(\varrho_0, \varrho_0 + d\varrho)$ nach Lemma 3b) die *Gleichung*

$$\sum_{\mu=1}^{m} \frac{1}{M_\mu(\varrho_0, \varrho_0 + d\varrho)} = \frac{1}{M(\varrho_0, \varrho_0 + d\varrho)}, \tag{7}$$

denn bei der konformen Abbildung von $G(\varrho_0, \varrho_0 + d\varrho)$ auf einen infinitesimalen Kreisring um den Nullpunkt der z-Ebene schwankt das Argument von z auf den Bildern der Schnittkurven σ beliebig wenig, wenn nur $d\varrho$ genügend klein ist, weil die Schnittkurven σ auf W stückweise analytisch sind und weil $B_{\varrho_0, w}$ jede Schnittkurve mit wohlbestimmter Tangente trifft, die mit ihr einen von Null verschiedenen Winkel bildet [nach einer quasikonformen Abbildung mit beschränktem Dilatationsquotienten, die σ fest läßt, durchsetzt $B_{\varrho_0, w}$ jede Schnittkurve im Kleinen spiegelbildlich].

Für ein $g^{0\zeta}$ in E_μ ist $\lambda_\mu(\varrho) = \frac{n}{\varrho}$, wenn am Rand des Bildes von g im Netz n Knoten liegen. Für ein $h_a^{0\zeta}$ ist $\lambda_\mu(\varrho) \equiv 1$. In einem $h_b^{0\zeta}$ oder einem $h_c^{0\zeta}$ liegt $\lambda_\mu(\varrho)$ zwischen 0 und 1 (die Grenzen ausgeschlossen). Konvergieren in einem Flächenstück h_b die Koordinaten der algebraischen Windungspunkte gegen die Koordinate des der zugehörigen Knotenfolge benachbarten logarithmischen Elementargebietes, so liegt $\lambda_\mu(\varrho)$ asymptotisch zwischen 0 und $\frac{1}{2}$; ein Grenzwert für $\lambda_\mu(\varrho)$ braucht selbstverständlich nicht zu existieren. Je nach der Stärke der Konvergenz kann $\lambda_\mu(\varrho)$ asymptotisch 0 (starke Konvergenz) oder $\frac{1}{2}$ (schwache Konvergenz) werden, jedoch ist stets $\lim \varrho\lambda_\mu(\varrho) = \infty$. Konvergieren dagegen die Koordinaten der algebraischen Windungspunkte gegen

eine andere als die oben erwähnte Koordinate [es liegt dann eine relativ mittelbare Randstelle über der Konvergenzkoordinate], so liegt $\lambda_\mu(\varrho)$ asymptotisch zwischen $\frac{1}{2}$ und 1 je nach der Stärke der Konvergenz; zum Wert $\frac{1}{2}$ gehört schwache, zum Wert 1 starke Konvergenz. Ganz analog liegen die Verhältnisse, wenn zu einem Flächenstück h_c genau eine (relativ mittelbare) Randstelle gehört; $\lambda_\mu(\varrho)$ liegt dann asymptotisch zwischen 0 und $\frac{1}{2}$ und es ist $\lim \varrho\,\lambda_\mu(\varrho) = \infty$. Gehören jedoch zu h_c zwei verschiedene (relativ mittelbare) Randstellen, so entspricht starker Konvergenz der Windungspunkte der asymptotische Wert 1 für $\lambda_\mu(\varrho)$, schwacher der asymptotische Wert 0. Wieder ist $\lim \varrho\,\lambda_\mu(\varrho) = \infty$. Hat aber kein Sinusende von h_c als Kern die punktierte reelle Achse, so gilt mit $a > 0$ die Beziehung $a\,\varrho^{-1} \leq \lambda_\mu(\varrho) = O(\varrho^{-\frac{1}{2}})$.

Wir setzen in (7) die durch (5), (6) gegebene Größe für $M_\mu(\varrho_0, \varrho_0 + d\varrho)$ ein und erhalten unter Beachtung von $M_\mu(\varrho_0, \varrho_0 + d\varrho) \geq c_\mu^{-1} m_\mu(\varrho_0, \varrho_0 + d\varrho)$, wo c_μ das Maximum des Dilatationsquotienten der Abbildung q_μ ist, nach kurzer Rechnung die Beziehung

$$(8)\quad \begin{cases} \dfrac{2\,d\varrho}{\sum\limits_{\mu=1}^{m} \lambda_\mu(\varrho_0)\cdot\varrho_0 + c(\varrho_0)} \leq M(\varrho_0, \varrho_0 + d\varrho) \\ \text{für ein } c(\varrho_0) \text{ mit } \varrho_0 \cdot \sum\limits_{\mu=1}^{m} \lambda_\mu(\varrho_0) > -c(\varrho_0) = O(1). \end{cases}$$

Nach Lemma 3a) erschließen wir für den Modul des Ringgebietes $G(\varrho', \varrho'')$ die Ungleichung

$$(9)\quad \int_{\varrho'}^{\varrho''} \frac{2}{\sum\limits_{\mu=1}^{m} \lambda_\mu(\varrho) + \dfrac{c(\varrho)}{\varrho}} \cdot \frac{d\varrho}{\varrho} \leq M(\varrho', \varrho'').$$

Da $M(\varrho', \varrho'')$ stets endlich, aber $\sum\limits_{\mu=1}^{m} \lambda_\mu(\varrho) \leq m$, strebt $M(\varrho', \varrho'')$ mit $\varrho'' \to \infty$ selbst gegen unendlich, es gilt also der

Satz 1. *Alle Flächen der Klasse* $\mathfrak{W}$ *haben parabolischen Typus*[25]).

6.2. Um mit Hilfe von Lemma 5 die Wertverteilung der meromorphen Funktion bestimmen zu können, die die Fläche W erzeugt, brauchen wir die zu (9) umgekehrte Ungleichung.

Es sei A die Menge derjenigen μ, für die Q_μ Bild eines Flächenstückes h^0 ist, B die Restmenge. Wir ziehen in E_μ, $\mu \in A$, zwei stückweise glatte Kurven b_μ^1, b_μ^2, die die folgenden Bedingungen erfüllen:

(10a) *sie laufen von* 0 *nach* ∞ *und sind, vom Endpunkt* ∞ *abgesehen, schließlich punktfremd*;

(10b) *sie enthalten schließlich nur Punkte des Innern oder des Randes von* $q_\mu(Q_\mu)$;

(10c) *sie durchsetzen jeden Kreis* $|\zeta_\mu| = \varrho$ *(nur einmal) unter von Null verschiedenem Winkel.*

[25]) Für ähnlich gebaute Flächen zeigen dies auf andere Weise bereits ULRICH [26] und MACLANE [11].

[Eine kleine Überlegung zeigt, daß die Bedingungen (10) wegen (F_1), (F_2) erfüllbar sind.]

Die Kurve b_μ^k, $k=1,2$, habe in Polarkoordinaten $\zeta_\mu = \varrho_\mu e^{i\varphi_\mu}$ die Gleichung $\varphi_\mu = \pi\lambda_\mu^k(\varrho)$; die Indizes seien so gewählt, daß $0 \leq \lambda_\mu^1(\varrho) < \lambda_\mu^2(\varrho) \leq 1$; das Stück $B_\varrho^{\mu *}$, das b_μ^1 und b_μ^2 aus B_ϱ^μ ausschneiden, hat die Länge $\pi\lambda_\mu^*(\varrho)\cdot\varrho$, wo $\lambda_\mu^* = \lambda_\mu^2 - \lambda_\mu^1$ ist. Der Modul $m_\mu^*(\varrho', \varrho'')$ des Vierecks $v_\mu^*(\varrho', \varrho'')$, das von b_μ^k als α-Bogen und von $B_{\varrho'}^{\mu *}$, $B_{\varrho''}^{\mu *}$ als β-Bogen begrenzt wird, ist (nach den Lemmata 1, 2) wenigstens ebenso groß wie $m_\mu(\varrho', \varrho'')$. Um $m_\mu^*(\varrho', \varrho'')$ nach oben abzuschätzen, bilden wir $v_\mu^*(\varrho', \varrho'')$ quasikonform ab in die ω_μ-Ebene, $\omega_\mu = \sigma_\mu \cdot e^{i\psi_\mu}$, durch

$$(11) \qquad \begin{cases} \sigma_\mu = \varrho_\mu \\ \psi_\mu = \dfrac{2}{\lambda_\mu^*(\varrho)}\left(\varphi_\mu - \pi\dfrac{\lambda_\mu^1 + \lambda_\mu^2}{2}\right). \end{cases}$$

Gleichzeitig werde $v_\mu^*(\varrho', \varrho'')$ konform durch $\log z_\mu(\zeta_\mu)$ so auf ein Rechteck abgebildet, daß die β-Seiten in Parallelseiten der Länge 2π übergehen; dann erhalten wir für den Modul $m_\mu^*(\varrho', \varrho'')$ die Beziehung

$$(12) \qquad m_\mu^*(\varrho', \varrho'') \leq \int_{\varrho'}^{\varrho''} \left|\frac{d\log z_\mu}{d\log\sigma_\mu}\right| d\log\sigma_\mu = \int_{\varrho'}^{\varrho''} \sqrt{\left|\frac{d\log z_\mu}{d\log\sigma_\mu}\right|^2}\, d\log\sigma_\mu.$$

Das Flächenelement in der $\log z_\mu$-Ebene bezeichnen wir mit $D\log z_\mu$, analog in anderen Ebenen ein Flächenelement durch Vorsetzen von D. Es ist dann

$$(13) \qquad \left|\frac{d\log z_\mu}{d\log\sigma_\mu}\right|^2 = \left|\frac{d\log z_\mu}{d\log\zeta_\mu}\right|^2 \cdot \left|\frac{d\log\zeta_\mu}{d\log\sigma_\mu}\right|^2 = \frac{D\log z_\mu}{D\log\zeta_\mu}\cdot\left|\frac{d\log\zeta_\mu}{d\log\sigma_\mu}\right|^2.$$

Mit

$$1 = \frac{D\log\zeta_\mu}{D\log\omega_\mu}\cdot\frac{D\log\omega_\mu}{D\log\zeta_\mu} = \frac{D\log\zeta_\mu}{D\log\omega_\mu}\cdot\frac{2}{\lambda_\mu^*}$$

erhalten wir

$$\left|\frac{d\log z_\mu}{d\log\sigma_\mu}\right|^2 = \frac{D\log z_\mu}{D\log\zeta_\mu}\cdot\frac{D\log\zeta_\mu}{D\log\omega_\mu}\cdot\frac{2}{\lambda_\mu^*}\left|\frac{d\log\zeta_\mu}{d\log\sigma_\mu}\right|^2.$$

Aus (11) gewinnen wir mit $|\psi_\mu| \leq \pi$

$$\begin{aligned}\left|\frac{d\log\zeta_\mu}{d\log\sigma_\mu}\right|^2 &= 1 + \frac{1}{4}\left\{\frac{d\lambda^*}{d\log\sigma_\mu}\psi_\mu + \pi\cdot\frac{d(\lambda_\mu^1+\lambda_\mu^2)}{d\log\sigma_\mu}\right\}^2 \\ &\leq 1 + \pi^2\cdot\left\{\left|\frac{d\lambda_\mu^1}{d\log\sigma_\mu}\right| + \left|\frac{d\lambda_\mu^2}{d\log\sigma_\mu}\right|\right\}^2;\end{aligned}$$

also ist

$$m_\mu^*(\varrho', \varrho'') \leq \int_{\varrho'}^{\varrho''} \sqrt{\frac{2}{\lambda_\mu^*}\left\{1 + \pi^2\left[\left|\frac{d\lambda_\mu^1}{d\log\sigma_\mu}\right| + \left|\frac{d\lambda_\mu^2}{d\log\sigma_\mu}\right|\right]^2\right\}}\cdot\sqrt{\frac{D\log z_\mu}{D\log\omega_\mu}}\cdot d\log\sigma_\mu.$$

Wir wenden die Schwarzsche Ungleichung an und erhalten

$$\{m_\mu^*(\varrho', \varrho'')\}^2 \leq \int_{\varrho'}^{\varrho''} \frac{2}{\lambda_\mu^*}\left\{1 + \pi^2\left[\left|\frac{d\lambda_\mu^1}{d\log\sigma_\mu}\right| + \left|\frac{d\lambda_\mu^2}{d\log\sigma_\mu}\right|\right]^2\right\} d\log\sigma_\mu \cdot \int_{\varrho'}^{\varrho''} \frac{D\log z_\mu}{D\log\omega_\mu}\, d\log\sigma_\mu.$$

Integration über ψ_μ von $-\pi$ bis $+\pi$ ergibt rechts den Faktor $2\pi\, m_\mu^*(\varrho', \varrho'')$. Division durch diesen Ausdruck gibt mit $\sigma = \varrho$ schließlich

$$(14) \qquad m_\mu^*(\varrho', \varrho'') \leq \int_{\varrho'}^{\varrho''} \frac{2}{\lambda_\mu^*(\varrho)} \left\{1 + \pi^2 \left[\left|\frac{d\lambda_\mu^1}{d\log\varrho}\right| + \left|\frac{d\lambda_\mu^2}{d\log\varrho}\right|\right]^2\right\} d\log\varrho.$$

Benutzen wir die Relation (14) für zwei Werte ϱ_1, ϱ_2 von ϱ, so erhalten wir

$$(15) \quad M_\mu(\varrho', \varrho'') \leq \int_{\varrho_1}^{\varrho_2} \frac{2}{\lambda_\mu(\varrho)} \frac{d\varrho}{\varrho} + (M_\mu(\varrho', \varrho'') - m_\mu^*(\varrho_1, \varrho_2)) + I_\mu(\varrho_1, \varrho_2) + J_\mu(\varrho_1, \varrho_2)$$

mit

$$I_\mu(\varrho_1, \varrho_2) = 2\pi^2 \int_{\varrho_1}^{\varrho_2} \frac{1}{\lambda_\mu^*(\varrho)} \left\{\left|\frac{d\lambda_\mu^1(\varrho)}{d\log\varrho}\right| + \left|\frac{d\lambda_\mu^2(\varrho)}{d\log\varrho}\right|\right\}^2 \frac{d\varrho}{\varrho}$$

und

$$J_\mu(\varrho_1, \varrho_2) = 2 \cdot \int_{\varrho_1}^{\varrho_2} \left(\frac{1}{\lambda_\mu^*(\varrho)} - \frac{1}{\lambda_\mu(\varrho)}\right) \frac{d\varrho}{\varrho}.$$

Aus (15) folgt nun mit (4) die Relation

$$(16) \quad \begin{cases} M(\varrho', \varrho'') \leq \left[\sum_{\mu=1}^{m} \{M_\mu(\varrho', \varrho'')\}^{-1}\right]^{-1} \\ \qquad\qquad \leq \left[\sum_{\mu \in A'} \{K_\mu(\varrho_1, \varrho_2) + M_\mu(\varrho', \varrho'') - m_\mu^*(\varrho_1, \varrho_2)\}^{-1}\right]^{-1}, \end{cases}$$

wobei A' eine Teilmenge von A ist, wenn wir noch

$$K_\mu(\varrho_1, \varrho_2) = \int_{\varrho_1}^{\varrho_2} \frac{2}{\lambda_\mu(\varrho)} \frac{d\varrho}{\varrho} + I_\mu(\varrho_1, \varrho_2) + J_\mu(\varrho_1, \varrho_2)$$

setzen. Für $\sum_{\mu=1}^{m} \lambda_\mu(\varrho) > c\varrho^{-\frac{1}{2}+\alpha}$ $(c > 0,\ \alpha > 0)$ ist nach (9)

$$(17) \qquad M(\varrho', \varrho'') - \int_{\varrho'}^{\varrho''} \frac{2}{\sum_{\mu=1}^{m} \lambda_\mu(\varrho)} \frac{d\varrho}{\varrho} \geq \varepsilon(\varrho').$$

Blicken wir auf Lemma 5 und vergleichen wir (16) mit (17), so erhebt sich die Frage: Unter welchen Voraussetzungen kann die Ungleichung (16) fortgesetzt werden durch

$$(16') \quad \left[\sum_{\mu \in A'} \{K_\mu(\varrho_1, \varrho_2) + M_\mu(\varrho', \varrho'') - m_\mu^*(\varrho_1, \varrho_2)\}^{-1}\right]^{-1} \leq \int_{\varrho'}^{\varrho''} \frac{2}{\sum_{\mu=1}^{m} \lambda_\mu(\varrho)} \frac{d\varrho}{\varrho} + \varepsilon(\varrho'),$$

wenigstens für $M(\varrho', \varrho'') \geq K > 0$?

6.3. Dies ist möglich in verschiedenen Fällen; von besonderer Bedeutung für die weiteren Untersuchungen ist der folgende. Es sei mit $\alpha > 0$

(18a) $\lambda_\mu(\varrho) = \Lambda_\mu + O(\varrho^{-\alpha})$ mit konstantem $\Lambda_\mu > 0$ für alle μ einer Menge $A_0 < A$;

(18b) $\lambda_\mu(\varrho) = O(\varrho^{-\alpha})$ für $\mu \in A_1 = A - A_0$;

(18c) $\sum_{\mu \in A_0} \Lambda_\mu > 0$.

Für den ν-ten Grätenendpunkt $z^\nu_{\mu 1} = \varrho^\nu_{\mu 1} e^{i\pi\lambda^\nu_{\mu 1}}$ bzw. $z^\nu_{\mu 2} = \varrho^\nu_{\mu 2} e^{i\pi\lambda^\nu_{\mu 2}}$ $(\nu = 1, 2, \ldots)$ des rechten bzw. linken Randes von $q_\mu(Q_\mu)$ auf $\Im(\zeta_\mu) = \nu\pi i$ gelte für $\mu \in A_0$

(18d) $$\lambda^\nu_{\mu k} - \tfrac{1}{2} = a_{\mu k}(\varrho^\nu_{\mu k})^{-\beta_{\mu k}} + O(1) \cdot (\varrho^\nu_{\mu k})^{-\beta'_{\mu k}}, \quad k = 1, 2,$$

wo $\beta'_{\mu k} > \beta_{\mu k} \geq 0$. Es ist wegen (18a) dann $\operatorname{Min}(\beta_{\mu 1}, \beta_{\mu 2}) = 0$. Wir wählen für b^k_μ folgende in Polarkoordinaten gegebene Kurve

$$\lambda^k_\mu(\varrho) - \tfrac{1}{2} = a_{\mu k}\varrho^{-\beta_{\mu k}} + a''_{\mu k}\varrho^{-\beta''_{\mu k}}$$

bei $\beta_{\mu k} < \beta''_{\mu k} < \beta'_{\mu k}$. Durch passende Wahl von $a''_{\mu k}$ werden die Bedingungen (10) erfüllt, überdies wird erreicht, daß (10b) erfüllt bleibt, wenn $q_\mu(Q_\mu)$ durch Q_μ ersetzt wird[26]).

Wir betrachten für $\mu \in A_0$ das quasikonforme q_μ^{-1}-Bild $V_\mu(\varrho', \varrho'')$ von $v_\mu(\varrho', \varrho'')$ und dazu das Gebiet $\widetilde{V}_\mu(\varrho', \varrho'')$, das aus Q_μ herausgehoben wird, durch diejenigen wohlbestimmten Kreisbogenstücke auf $|\zeta_\mu| = \varrho'$ und $|\zeta_\mu| = \varrho''$, die den linken und den rechten Rand von Q_μ verbinden. $V_\mu(\varrho', \varrho'')$ entsteht nach (Q_3), (3) aus $v_\mu(\varrho', \varrho'')$ dadurch, daß die α-Mengen von $v_\mu(\varrho', \varrho'')$, also die Gräten, auf den Geraden $\Im(\zeta_\mu) = \nu\pi i$ um $O(1)$, und die β-Mengen, also die Kreisbogenstücke $B^\mu_{\varrho'}$, $B^\mu_{\varrho''}$, punktweise um $O(1)$ verschoben werden — natürlich so, daß sie zusammenhängend bleiben. Jedenfalls geht aber, unabhängig von ϱ, für hinreichend großes $c > 0$ für alle $\mu \in A_0$ das Gebiet $\widetilde{V}_\mu(\varrho' - c, \varrho'' + c)$ aus $V_\mu(\varrho', \varrho'')$ hervor durch Gebietserweiterung über β-Bogen und nachträgliches Ersetzen von β-Mengen durch α-Mengen. Nach den Lemmata 1, 2 ist also der Modul $\widetilde{M}_\mu(\varrho' - c, \varrho'' + c)$ von $\widetilde{V}_\mu(\varrho' - c, \varrho'' + c)$ größer als $M_\mu(\varrho', \varrho'')$. Weil b^k_μ, $k = 1, 2$, auch innerhalb Q_μ liegt, ist (ebenfalls nach den Lemmata 1, 2) auch

(19) $$M_\mu(\varrho', \varrho'') \leq \widetilde{M}_\mu(\varrho' - c, \varrho'' + c) \leq m^*_\mu(\varrho' - c, \varrho'' + c).$$

Es gelten für $\mu \in A_0$ mit geeignetem β, $0 < \beta < \alpha$, nach der Wahl von b^k_μ die Beziehungen $\lambda^*_\mu(\varrho) = \Lambda_\mu + O(\varrho^{-\beta})$ und $\frac{d\lambda^k_\mu}{d\log\varrho} = O(\varrho^{-\beta})$. Daher ist

$$I_\mu(\varrho' - c, \varrho'' + c) = O(1) \cdot \int_{\varrho'-c}^{\varrho''+c} \varrho^{-2\beta} \cdot \frac{d\varrho}{\varrho} = O(1) \cdot (\varrho' - c)^{-2\beta}\left[1 - \left(\frac{\varrho'' + c}{\varrho' - c}\right)^{-2\beta}\right]$$

und

$$J_\mu(\varrho' - c, \varrho'' + c) = O(1) \cdot \int_{\varrho'-c}^{\varrho''+c} \varrho^{-\beta} \frac{d\varrho}{\varrho} = O(1) \cdot (\varrho' - c)^{-\beta}\left[1 - \left(\frac{\varrho'' + c}{\varrho' - c}\right)^{-\beta}\right].$$

[26]) Man beachte $\mu \in A_0$ und (F_2).

Es ist daher für diese μ

$$(20)\quad K_\mu(\varrho'-c,\varrho''+c)=\frac{2}{\Lambda_\mu}\cdot\log\frac{\varrho''+c}{\varrho'-c}+O(1)\cdot(\varrho'-c)^{-\beta}\left[1-\left(\frac{\varrho''+c}{\varrho'-c}\right)^{-\beta}\right].$$

Nach (16) ist also mit $\varrho_1=\varrho'-c$, $\varrho_2=\varrho''+c$ wegen (19)

$$M(\varrho',\varrho'')\leq\Big[\sum_{\mu\in A_0}\{K_\mu(\varrho'-c,\varrho''+c)\}^{-1}\Big]^{-1}.$$

Hieraus folgt: Notwendig dafür, daß $M(\varrho',\varrho'')\geq K>0$ für hinreichend großes ϱ', ist nach (20) $\frac{\varrho''}{\varrho'}=L\geq L_0$ für ein geeignetes $L_0>1$. Mit $\Lambda=\sum_{\mu\in A_0}\Lambda_\mu$ erhalten wir so für $L\geq L_0$

$$(21)\qquad M(\varrho',\varrho'')\leq\frac{2}{\Lambda}\log L+O(\varrho'^{-\beta}).$$

Umgekehrt ist

$$(21')\qquad \int_{\varrho'}^{\varrho''}\frac{2}{\sum_{\mu=1}^{m}\lambda_\mu(\varrho)}\frac{d\varrho}{\varrho}=\int_{\varrho'}^{\varrho''}\frac{2}{\Lambda+O(\varrho^{-\alpha})}\frac{d\varrho}{\varrho}=\frac{2}{\Lambda}\log L+O(\varrho'^{-\alpha}),$$

und aus (21), (21') zusammen folgt die Richtigkeit der Fortsetzung (16') von (16) unter den Bedingungen (18).

6.4. Wir nehmen an, es sei $\sum_{\mu=1}^{m}\lambda_\mu(\varrho)>c\varrho^{-\frac{1}{2}+\alpha}$ $(c>0,\ \alpha>0)$, und es gelte (16') als Fortsetzung von (16) mit geeigneten ϱ_1, ϱ_2 sobald $M(\varrho',\varrho'')\geq K>0$ und ϱ', ϱ'' hinreichend groß; es bestehe also für diese Werte ϱ', ϱ'' die Beziehung

$$(22)\qquad M(\varrho',\varrho'')-\int_{\varrho'}^{\varrho''}\frac{2}{\sum_{\mu=1}^{m}\lambda_\mu(\varrho)}\frac{d\varrho}{\varrho}=\varepsilon(\varrho').$$

Sei $w(z)$ die analytische Funktion, die W als Bild der punktierten Ebene[27] $|z|<\infty$ erzeugt. $B_{\varrho,z}$ sei in der z-Ebene das Urbild der geschlossenen Kurve $B_{\varrho,w}$. Nach Lemma 5 erschließen wir mit einer Konstanten $C>0$ aus (22) die Beziehung

$$(23)\quad u_j(\varrho)=Ce^{f(\varrho)}(1+\varepsilon(\varrho))\quad\text{mit}\quad f(\varrho)=\int_{\varrho_0}^{\varrho}\frac{2}{\sum_{\mu=1}^{m}\lambda_\mu(\varrho)}\frac{d\varrho}{\varrho},\quad j=1,2,$$

wo $u_1(\varrho)=\operatorname{Min}_{B_{\varrho,z}}|z|$, $u_2(\varrho)=\operatorname{Max}_{B_{\varrho,z}}|z|$ und ϱ_0 eine feste genügend große Zahl ist.

Die Funktion $u_j(\varrho)$ ist eigentlich monoton und stückweise stetig [sie ist möglicherweise unstetig an einer Unstetigkeitsstelle eines $\lambda_\mu(\varrho)$]. Wir definieren daher neben $u_j(\varrho)$ die Funktion $\bar{u}_j(\varrho)$: sie soll mit $u_j(\varrho)$ bis auf die Sprungstellen übereinstimmen, an diesen soll sie jedoch alle Werte zwischen

[27] Vgl. Satz 1.

den Sprungordinaten, diese eingeschlossen, annehmen. Die (eindeutige) Umkehrfunktion $\varrho_j(r)$ von $\bar{u}_j(\varrho)$ bezeichnen wir kurz als *Umkehrfunktion von* $u_j(\varrho)$. $\varrho_j(r)$ ist insbesondere stetig. Es ist

$$u_j(\varrho_j(r)-0) \leq r \leq u_j(\varrho_j(r)+0). \tag{24}$$

Aus (23) folgt nun

$$f(\varrho_1(r)) - f(\varrho_2(r)) = \varepsilon(r). \tag{25}$$

Wir bemerken ferner [wegen $\lambda_\mu \leq 1$]

$$df(\varrho) \geq \frac{2}{m}\, d\log\varrho; \tag{26}$$

aus (25) folgt also

$$\frac{\varrho_1(r)}{\varrho_2(r)} = 1 + \varepsilon(r). \tag{27}$$

6.5. Man bezeichnet mit $n(r, a)$ die Anzahl der mit ihrer Vielfachheit gezählten a-Stellen von $w(z)$ im Kreis $|z|<r$; $n_1(r, a)$ zählt eine λ-fache a-Stelle nur $\lambda-1$ mal. Die Anzahl der a-Stellen innerhalb $B_{\varrho,z}$ wollen wir mit $\nu(\varrho, a)$ bezeichnen und entsprechend die λ-fachen a-Stellen innerhalb dieser Kurve $\lambda-1$ mal mit $\nu_1(\varrho, a)$ zählen. Die bezüglich $B_{\varrho,z}$ bestimmten Anzahlfunktionen entnimmt man aus der Anzahl der verschiedenen [in Q_μ auf Parallelen zur imaginären Achse im Abstand 2π liegenden] Werte $\zeta^{(\mu)}(a)$ in den einzelnen Halbebenen E_μ, die nach der Transformation q_μ innerhalb B_ϱ^μ liegen, d.h. durch B_ϱ^μ von den in $q_\mu(Q_\mu)$ auf der Geraden $\Im(\zeta_\mu)=[\varrho+1]$ liegenden Punkten getrennt werden [Werte von $q_\mu\zeta^{(\mu)}(a)$ auf dem Rand von $q_\mu(Q_\mu)$ sind dabei nur halb so oft zu zählen]. Es sei nämlich $\nu^{(\mu)}(\varrho, a)$ bzw. $\nu_1^{(\mu)}(\varrho, a)$ die Anzahl der λ-fach bzw. $\overline{\lambda-1}$-fach gezählten a-Stellen mit Bildern in $q_\mu(Q_\mu)$ innerhalb B_ϱ^μ, $\tilde{\nu}^{(\mu)}(\varrho, a)$ bzw. $\tilde{\nu}_1^{(\mu)}(\varrho, a)$ dieselbe Anzahl mit $\zeta^{(\mu)}$-Bildern in Q_μ innerhalb $\tilde{B}_\varrho^\mu$. Aus (Q_3) [für $f_\mu \in H$], (Q_4) [für $f_\mu \in G$] und (3) erschließt man die Existenz einer Zahl $c>0$, so daß für alle μ die Beziehungen

$$\begin{cases} \tilde{\nu}^{(\mu)}(\varrho+c, a) \geq \nu^{(\mu)}(\varrho, a), & \nu^{(\mu)}(\varrho+c, a) \geq \tilde{\nu}^{(\mu)}(\varrho, a) \\ \tilde{\nu}_1^{(\mu)}(\varrho+c, a) \geq \nu_1^{(\mu)}(\varrho, a), & \nu_1^{(\mu)}(\varrho+c, a) \geq \tilde{\nu}_1^{(\mu)}(\varrho, a) \end{cases} \tag{28}$$

gelten [vgl. die analoge Überlegung auf S. 259]. Es sei $\nu^{(\mu)}(\varrho)$ die halbe Anzahl der Streifen $(\nu-1)\pi \leq \Im(\zeta_\mu) \leq \nu\pi$ $(\nu=1, 2, \ldots)$, die schon Punkte innerhalb B_ϱ^μ enthalten; dann ist für nicht-reelles a_0

$$\nu^{(\mu)}(\varrho) = \nu^{(\mu)}(\varrho, a_0) + O(1).$$

Weil $\nu^{(\mu)}(\varrho, a) = O(\varrho)$ gleichmäßig in a gilt, ist daher

$$\nu(\varrho) = \sum_{\mu=1}^{m} \nu^{(\mu)}(\varrho) = \operatorname*{Max}_a \nu(\varrho, a) + O(1) = O(\varrho). \tag{29}$$

Die nach $\log r$ bzw. $f(\varrho)$ integrierten Anzahlen werden mit dem entsprechenden Großbuchstaben bezeichnet, so daß mit festen r_0, ϱ_0 gilt

$$N(r,a)=\int_{r_0}^{r} n(r,a)\,\frac{dr}{r}, \qquad N_1(r,a)=\int_{r_0}^{r} n_1(r,a)\,\frac{dr}{r}$$

$$\mathbf{N}(\varrho,a)=\int_{\varrho_0}^{\varrho} \nu(\varrho,a)\,df(\varrho), \qquad \mathbf{N}_1(\varrho,a)=\int_{\varrho_0}^{\varrho} \nu_1(\varrho,a)\,df(\varrho).$$

Wir setzen ferner

$$N(r)=\max_a N(r,a), \qquad \mathbf{N}(\varrho)=\int_{\varrho_0}^{\varrho} \nu(\varrho)\,df(\varrho).$$

Zwischen den „lateinischen" und den „griechischen" Anzahlen bestehen nach der Definition der Umkehrfunktionen $\varrho_j(r)$ und $u_j(\varrho)$ die Beziehungen

$$\begin{cases} \nu\left(\varrho_2(r)-0,a\right)\leq n(r,a)\leq \nu\left(\varrho_1(r)+0,a\right) \\ \nu_1\left(\varrho_2(r)-0,a\right)\leq n_1(r,a)\leq \nu_1\left(\varrho_1(r)+0,a\right). \end{cases} \tag{30}$$

Wir fordern, es sei für jedes $K>0$ und für ein $\gamma>0$

$$\lim \frac{\nu(\varrho+K)}{\nu(\varrho)}=1, \qquad \overline{\lim}\,\frac{\nu(\varrho+\gamma\varrho)}{\nu(\varrho)}<\infty; \tag{31}$$

[hieraus folgt insbesondere

$$\overline{\lim}\,\frac{\nu(\varrho)}{\mathbf{N}(\varrho)}<\infty.] \tag{32}$$

Dann können Ordnung, Typus, die Defekte und die Indizes der algebraischen Verzweigtheit für die Funktion $w(z)$ sogleich bestimmt werden.

6.6. Als (obere) Wachstumsordnung s^* bzw. als untere Wachstumsordnung s_* bezeichnet man die Limites

$$s^*=\overline{\lim}\,\frac{\log T(r,w)}{\log r}, \qquad s_*=\underline{\lim}\,\frac{\log T(r,w)}{\log r};$$

$T(r,w)$ ist dabei die Charakteristik der meromorphen Funktion $w(z)$. Der obere bzw. untere Wachstumstypus der oberen bzw. unteren Ordnung heißt maximal, normal (mittel) oder minimal, je nachdem

$$\overline{\lim}\,\frac{T(r,w)}{r^{s^*}} \quad \text{bzw.} \quad \underline{\lim}\,\frac{T(r,w)}{r^{s_*}}$$

unendlich, positiv endlich oder Null ist. Es ist hier das Wachstum der Charakteristik $T(r,w)$ gleich dem von $N(r)$; denn nach dem Satz von COLLINGWOOD-TEICHMÜLLER[28]) gilt hier für jede nicht-reelle Zahl a_0 die Gleichung

$$T(r,w)=N(r,a_0)+O(1)=N(r)+O(1).$$

Der Defekt $\delta(a)$ eines Wertes a ist definiert durch

$$\delta(a)=1-\overline{\lim}\,\frac{N(r,a)}{T(r,w)},$$

[28]) Siehe z. B. TEICHMÜLLER [18].

der Index der algebraischen Verzweigtheit $\vartheta(a)$ durch

$$\vartheta(a) = \underline{\lim} \frac{N_1(r, a)}{T(r, w)}.$$

Mit der Beziehung (24), [in der für gewisse r links und für gewisse r rechts Gleichheit herrscht] und mit (30) erhalten wir wegen der Monotonie von $\nu(\varrho, a)$ für nichtreelles a_0

$$O(1) + \int_{\varrho_0}^{\varrho_2(r)-0} \nu(\varrho - 0, a_0)\, d \log u_2(\varrho) \leq \int_{r_0}^{r} n(r, a_0)\, d \log r$$

$$= N(r, a_0) \leq \int_{\varrho_0}^{\varrho_1(r)+0} \nu(\varrho + 0, a_0)\, d \log u_1(\varrho) + O(1)$$

und hieraus nach (23)

$$\varepsilon(r) \cdot \nu\left(\varrho_2(r) - 0, a_0\right) + \int_{\varrho_0}^{\varrho_2(r)} \nu(\varrho, a_0)\, d f(\varrho) \leq N(r, a_0)$$

$$= N(r) + O(1) \leq \int_{\varrho_0}^{\varrho_1(r)} \nu(\varrho, a_0)\, d f(\varrho) + \varepsilon(r)\, \nu\left(\varrho_1(r) + 0, a_0\right).$$

Mit (25), (27), (29), (31), (32) folgt nun

$$N(r) = \mathsf{N}\left(\varrho_1(r)\right)\left[1 + \varepsilon(r)\right] = \mathsf{N}\left(\varrho_2(r)\right)\left[1 + \varepsilon(r)\right], \tag{33}$$

mit (24) ist also

$$\frac{\log \mathsf{N}\left(\varrho_1(r)\right) + \varepsilon(r)}{\log u_1\left(\varrho_1(r) + 0\right)} \leq \frac{\log N(r)}{\log r} \leq \frac{\log \mathsf{N}\left(\varrho_2(r)\right) + \varepsilon(r)}{\log u_2\left(\varrho_2(r) - 0\right)}.$$

Nach (23) folgt

$$\overline{\lim} \frac{\log \mathsf{N}(\varrho)}{f(\varrho)} = \overline{\lim} \frac{\log N(r)}{\log r} = s^*, \tag{34*}$$

$$\underline{\lim} \frac{\log \mathsf{N}(\varrho)}{f(\varrho)} = \underline{\lim} \frac{\log N(r)}{\log r} = s_*. \tag{34$_*$}$$

Man schließt genau so (mit reellem s)

$$\frac{\mathsf{N}\left(\varrho_1(r)\right)\left[1 + \varepsilon(r)\right]}{\left[u_1\left(\varrho_1(r) + 0\right)\right]^s} \leq \frac{N(r)}{r^s} \leq \frac{\mathsf{N}\left(\varrho_2(r)\right)\left[1 + \varepsilon(r)\right]}{\left[u_2\left(\varrho_2(r) - 0\right)\right]^s},$$

und daraus auf den oberen bzw. unteren Maximal-, Normal- (Mittel-) oder Minimaltypus der oberen bzw. unteren Ordnung, je nachdem die Limites

$$\overline{\lim} \frac{\mathsf{N}(\varrho)}{e^{s^* f(\varrho)}} = \overline{\lim} \frac{N(r)}{r^{s^*}} \cdot C^{s^*} \tag{35*}$$

bzw.

$$\underline{\lim} \frac{\mathsf{N}(\varrho)}{e^{s_* f(\varrho)}} = \underline{\lim} \frac{N(r)}{r^{s_*}} \cdot C^{s_*} \tag{35$_*$}$$

unendlich, positiv endlich oder Null sind. Gilt über (31) hinaus noch für jedes $K > 0$

$$\lim \frac{\nu(\varrho + K, a)}{\nu(\varrho, a)} = 1 \quad \text{bzw.} \quad \lim \frac{\nu_1(\varrho + K, a)}{\nu_1(\varrho, a)} = 1 \tag{36}$$

und für ein $\gamma > 0$

$$(36') \qquad \overline{\lim} \frac{\nu(\varrho+\gamma\varrho, a)}{\nu(\varrho, a)} < \infty \quad \text{bzw.} \quad \overline{\lim} \frac{\nu_1(\varrho+\gamma\varrho, a)}{\nu_1(\varrho, a)} < \infty,$$

so folgt ganz ähnlich

$$\frac{N(r, a)}{N(r)} = \frac{\mathsf{N}(\varrho_1(r), a)}{\mathsf{N}(\varrho_1(r))}\left(1+\varepsilon(r)\right) = \frac{\mathsf{N}(\varrho_2(r), a)}{\mathsf{N}(\varrho_2(r))}\left(1+\varepsilon(r)\right)$$

bzw.

$$\frac{N_1(r, a)}{N(r)} = \frac{\mathsf{N}_1(\varrho_1(r), a)}{\mathsf{N}(\varrho_1(r))}\left(1+\varepsilon(r)\right) = \frac{\mathsf{N}_1(\varrho_2(r), a)}{\mathsf{N}(\varrho_2(r))}\left(1+\varepsilon(r)\right),$$

und hieraus ergibt sich für den Defekt $\delta(a)$ bzw. den Verzweigungsindex $\vartheta(a)$ die Formel

$$(37) \qquad \delta(a) = 1 - \overline{\lim} \frac{\mathsf{N}(\varrho, a)}{\mathsf{N}(\varrho)} \quad \text{bzw.} \quad \vartheta(a) = \underline{\lim} \frac{\mathsf{N}_1(\varrho, a)}{\mathsf{N}(\varrho)}.$$

Mit

$$\begin{aligned} N(r, a) &\leq \int_{\varrho_0}^{\varrho_1(r)+0} \nu(\varrho+0, a)\, d\log u_1(\varrho) + O(1) \\ &= \mathsf{N}\left(\varrho_1(r), a\right) + \varepsilon(r)\left[\mathsf{N}\left(\varrho_1(r), a\right) + \nu\left(\varrho_1(r)+0, a\right)\right], \\ N(r, a) &\geq \int_{\varrho_0}^{\varrho_2(r)-0} \nu(\varrho-0, a)\, d\log u_2(\varrho) + O(1) \\ &= \mathsf{N}\left(\varrho_2(r), a\right) + \varepsilon(r)\left[N\left(\varrho_2(r), a\right) + \nu\left(\varrho_2(r)-0, a\right)\right] \end{aligned}$$

und der ähnlich gewonnenen Ungleichung

$$\begin{aligned} &\mathsf{N}_1\left(\varrho_2(r), a\right) + \varepsilon(r)\left[\mathsf{N}_1\left(\varrho_2(r), a\right) + \nu\left(\varrho_2(r)-0, a\right)\right] \\ &\qquad \leq N_1(r, a) \leq \mathsf{N}_1\left(\varrho_1(r)\right) + \varepsilon(r)\left[\mathsf{N}\left(\varrho_1(r), a\right) + \nu\left(\varrho_1(r)+0, a\right)\right] \end{aligned}$$

erschließt man aus (33) aber jedenfalls [ohne Verwendung von (36), (36')] die Gleichungen

$$(38) \qquad \begin{cases} \delta(a) = 1 - \overline{\lim} \dfrac{\mathsf{N}(\varrho, a) + \varepsilon(\varrho)\cdot\nu(\varrho+0, a)}{\mathsf{N}(\varrho)} \\ \text{und} \\ \vartheta(a) = \underline{\lim} \dfrac{\mathsf{N}_1(\varrho, a) + \varepsilon(\varrho)\cdot\nu_1(\varrho+0, a)}{\mathsf{N}(\varrho)}. \end{cases}$$

6.7. Wir fassen — unter Verwendung der Tatsache, daß $\nu^{(\mu)}(\varrho) = O(\varrho)$, aber daß für $f_\mu \in G$ jedoch $\nu^{(\mu)}(\varrho) = O(1)$ und $\lambda_\mu(\varrho) = O\left(\frac{1}{\varrho}\right)$ ist, und daß alle mehrfachen a-Stellen genau doppelte a-Stellen sind — die Ergebnisse zusammen im

Satz 2. *Es erzeuge $w(z)$ eine* RIEMANNsche *Fläche $W \in \mathfrak{W}$. Die Fläche werde nach § 3 zerschnitten und jedes ausfallende Flächenstück $f_{\mu_\tau} \in H$ ($\tau = 1, 2, \ldots, t$) werde durch $\zeta^{(\mu_\tau)} = \log\dfrac{w - w'_{\mu_\tau}}{w - w''_{\mu_\tau}} + i\,\gamma^{(\mu_\tau)}$ in die Grätenfigur Q_{μ_τ} der oberen Halbebene $\Im(\zeta_{\mu_\tau}) > 0$ abgebildet; hierbei sei w'_{μ_τ} bzw. w''_{μ_τ} die Koordinate*

des linken bzw. rechten Nachbargebietes der zu f_{μ_τ} gehörigen Knotenfolge im Streckenkomplex. Mit den durch (Q_1) bis (Q_4) bestimmten Transformationen q_{μ_τ} sei $B_\varrho^{\mu_\tau}$ das Stück des Kreisbogens $|\zeta_{\mu_\tau}| = \varrho$, das den rechten und den linken Rand von $q_{\mu_\tau}(Q_{\mu_\tau})$ verbindet, die Länge von $B_\varrho^{\mu_\tau}$ sei $\pi \lambda_{\mu_\tau}(\varrho) \cdot \varrho$. Es sei $\nu^{(\mu_\tau)}(\varrho)$ die halbe Anzahl der Streifen $(\nu-1)\pi \leq \Im(\zeta_{\mu_\tau}) \leq \nu\pi$ $(\nu = 1, 2, \ldots)$, die schon Punkte innerhalb $B_\varrho^{\mu_\tau}$ enthalten; es werde gesetzt

$$(39) \qquad f(\varrho) = \int_{\varrho_0}^{\varrho} \frac{2}{\sum_{\tau=1}^{t} \lambda_{\mu_\tau}(\varrho)} \frac{d\varrho}{\varrho}, \quad \nu(\varrho) = \sum_{\tau=1}^{t} \nu^{(\mu_\tau)}(\varrho), \quad \mathsf{N}(\varrho) = \int_{\varrho_0}^{\varrho} \nu(\varrho)\, df(\varrho).$$

a) *Wenn dann* $\sum_{\tau=1}^{t} \lambda_{\mu_\tau}(\varrho) > c\varrho^{-\frac{1}{2}+\alpha}$ *ist, mit* $c > 0$ *und* $\alpha > 0$, *wenn ferner für jedes* $K > 0$ *und ein* $\gamma > 0$ *gilt*

$$(31) \qquad \lim \frac{\nu(\varrho + K)}{\nu(\varrho)} = 1, \quad \overline{\lim} \frac{\nu(\varrho + \gamma\varrho)}{\nu(\varrho)} < \infty,$$

und wenn die Ungleichung (16) *durch* (16′) *fortgesetzt werden kann, so sind obere bzw. untere Ordnung von* $w(z)$ *gegeben durch*

$$s^* = \overline{\lim} \frac{\log \mathsf{N}(\varrho)}{f(\varrho)} \quad bzw. \quad s_* = \underline{\lim} \frac{\log \mathsf{N}(\varrho)}{f(\varrho)}.$$

Der obere bzw. untere Wachstumstypus zur oberen Ordnung s^* *bzw. zur unteren Ordnung* s_* *ist maximal, normal (mittel) oder minimal, je nachdem der Wert von*

$$\overline{\lim} \frac{\mathsf{N}(\varrho)}{e^{s^* f(\varrho)}} \quad bzw. \quad \underline{\lim} \frac{\mathsf{N}(\varrho)}{e^{s_* f(\varrho)}}$$

unendlich, positiv endlich oder Null ist.

b) *Bezeichnen wir dazu mit* $\nu(\varrho, a)$ *bzw.* $\nu_1(\varrho, a)$ *die Gesamtanzahl der a-Stellen überhaupt bzw. der mehrfachen a-Stellen, deren* $q_{\mu_\tau}\zeta^{(\mu_\tau)}$*-Bilder in irgendeinem* $q_{\mu_\tau}(Q_{\mu_\tau})$ *schon innerhalb* $B_\varrho^{(\mu_\tau)}$ *liegen [wobei Bilder auf dem Rand eines* $q_{\mu_\tau}(Q_{\mu_\tau})$ *nur halb zu zählen sind] und gilt außerdem für alle* $K > 0$ *und ein geeignetes* $\gamma > 0$

$$(36) \qquad \lim \frac{\nu(\varrho + K, a)}{\nu(\varrho, a)} = 1 \quad bzw. \quad \lim \frac{\nu_1(\varrho + K, a)}{\nu(\varrho)} = 1$$

sowie

$$(36') \qquad \overline{\lim} \frac{\nu(\varrho + \gamma\varrho, a)}{\nu(\varrho, a)} < \infty \quad bzw. \quad \overline{\lim} \frac{\nu_1(\varrho + \gamma\varrho, a)}{\nu_1(\varrho, a)} < \infty,$$

so ist der Defekt $\delta(a)$ *und der Index der algebraischen Verzweigtheit* $\vartheta(a)$ *gegeben durch*

$$(37) \qquad \delta(a) = 1 - \overline{\lim} \frac{\mathsf{N}(\varrho, a)}{\mathsf{N}(\varrho)} \quad bzw. \quad \vartheta(a) = \underline{\lim} \frac{\mathsf{N}_1(\varrho, a)}{\mathsf{N}(\varrho)},$$

wo

$$\mathsf{N}(\varrho, a) = \int_{\varrho_0}^{\varrho} \nu(\varrho, a)\, df(\varrho) \quad und \quad \mathsf{N}_1(\varrho, a) = \int_{\varrho_0}^{\varrho} \nu_1(\varrho, a)\, df(\varrho).$$

Jedenfalls aber gelten die Gleichungen

$$(38)\quad \begin{cases} \delta(a) = 1 - \overline{\lim}\, \dfrac{N(\varrho, a) + \varepsilon(\varrho)\cdot\nu(\varrho+0, a)}{N(\varrho)} \\ \text{und} \\ \vartheta(a) = \underline{\lim}\, \dfrac{N_1(\varrho, a) + \varepsilon(\varrho)\cdot\nu_1(\varrho+0, a)}{N(\varrho)}. \end{cases}$$

Satz 2 erlaubt, die Wertverteilung der Funktion $w(z)$ allein aus der Struktur der erzeugten RIEMANNschen Fläche zu entnehmen; wir bemerken, daß die Voraussetzungen des Satzes 2a) für die Flächen erfüllt sind, die den Bedingungen (18) genügen. Der Satz kann in folgender prägnanter Form ausgesprochen werden: *Die Voraussetzungen des Satzes erlauben, die „lateinischen" Anzahlfunktionen durch die „griechischen" zu ersetzen.*

Neben Satz 2 besteht

Satz 3. *Gilt mit den Bezeichnungen von Satz 2 die Gleichung* $\lim \sum_{\tau=1}^{t} \lambda_{\mu_\tau}(\varrho) = 0$, *so wird W von einer meromorphen Funktion der Ordnung Null erzeugt.*

Satz 3 wird in § 9 bewiesen.

§ 7. Die Stärke von Randstellen

Eine in einem Sinusende E einer RIEMANNschen Fläche vorliegende Konvergenz der Koordinaten der Windungspunkte beeinflußt das Wachstum der erzeugenden Funktion in entgegengesetzter Weise, je nachdem der Hüllenkern $kh(E)$ des Endes aus einem Punkt oder aus der punktierten reellen Achse besteht. Im ersten Fall trägt E eine (relativ mittelbare) Randstelle[29]), wir nennen E daher ein *randstelletragendes Ende*; im anderen Fall verschmelzen die beiden den benachbarten Randgebieten zugehörigen Relativ-Randstellen (die — zunächst einmal vorhanden sind und — notwendig über derselben Koordinate liegen) zu einer einzigen Randstelle[30]), wir nennen E dann ein *verschmelzendes Ende*. Wie sich zeigen wird, hat ein Verschmelzen von Randstellen die Tendenz, das Wachstum zu mindern, während das Auftreten von neuen Randstellen eher das Wachstum fördert. Ein Beispiel für verschmelzende und für randstelletragende Sinusenden gibt Fig. 3.

In diesem Paragraphen bedeute α eine Zahl > 0, aber nicht notwendig immer dieselbe. Um die Stärke einer Randstelle über a zu definieren, normieren wir auf $a = \infty$ als Konvergenzkoordinate eines verschmelzenden oder eines randstelletragenden Sinusendes; das bedeutet offensichtlich nur eine Vereinfachung der Schreibarbeit.

Definition 1. Ein randstelletragendes Sinusende einer Fläche $W \in \mathfrak{W}$, in dem die Koordinaten a_n der Windungspunkte gegen ∞ konvergieren, trägt

[29]) Dies zeigt auch Satz 2 in [7].
[30]) Dies zeigt auch Satz 3 in [7].

eine (relativ mittelbare) Randstelle der *Konvergenz-* oder *Relativstärke* $p \geq 0$, wenn gilt

$$\log|a_n| = pn + O(n^{1-\alpha}), \tag{40}$$

und der *Stärke* ∞, wenn gilt

$$n^{1+\alpha} = O(\log|a_n|). \tag{40'}$$

Definition 2. Eine Relativ-Randstelle, die einem logarithmischen Elementargebiet entspricht, hat die Konvergenz- oder Relativstärke ∞.

Definition 3. Ein verschmelzendes Sinusende einer Fläche $W \in \mathfrak{W}$ in dem die Koordinaten a_n der Windungspunkte gegen ∞ konvergieren, hat die *Verschmelzungsstärke* $p \geq 0$, wenn gilt

$$\log|a_n| = pn + O(n^{1-\alpha}), \tag{41}$$

es hat die *Verschmelzungsstärke* ∞, wenn gilt

$$n^{1+\alpha} = O(\log|a_n|). \tag{41'}$$

Definition 4. Ein Sinusende der Koordinate ∞, in dem die Koordinaten der Windungspunkte nicht konvergieren, während aber (40) mit $p = 0$ gilt, werde als *Ende der Stärke Null* bezeichnet.

Wir betrachten eine Halbebene E_μ, in die ein Flächenstück h_b^0 durch $\zeta^{(\mu)}(w)$ abgebildet sei, h_b habe das Sinusende im linken Teil. Ist E ein randstelletragendes Ende der Stärke p $(0 \leq p \leq \infty)$, so ist die Koordinate z_μ^n des n-ten Grätenendpunktes gegeben durch

$$\begin{cases} z_\mu^n = -pn + O(n^{1-\alpha}) + n\pi i & \text{für} \quad p \neq \infty, \\ z_\mu^n = -\gamma_n \cdot n^{1+\alpha} + n\pi i & \text{mit} \quad \underline{\lim}\, \gamma_n > 0 \quad \text{für} \quad p = \infty. \end{cases} \tag{42}$$

Ist E ein verschmelzendes Ende der Stärke p, so ist

$$\begin{cases} z_\mu^n = pn + O(n^{1-\alpha}) + n\pi i & \text{für} \quad p \neq \infty \\ z_\mu^n = \gamma_n \cdot n^{1+\alpha} + n\pi i & \text{mit} \quad \underline{\lim}\, \gamma_n > 0 \quad \text{für} \quad p = \infty. \end{cases} \tag{42'}$$

Die Gln. (42), (42') gelten nach (Q_3), (3) sowohl vor wie nach der quasikonformen Abbildung q_μ. Demgemäß gilt für $\lambda_\mu(\varrho)$ die Beziehung

$$\lambda_\mu(\varrho) = \frac{1}{2} + \frac{1}{\pi} \operatorname{arc\,tg} \frac{p}{\pi} + O(\varrho^{-\alpha}) \quad \text{beim randstelletragenden Sinusende,} \tag{43}$$

$$\lambda_\mu(\varrho) = \frac{1}{\pi} \operatorname{arc\,tg} \frac{\pi}{p} + O(\varrho^{-\alpha}) \quad \text{beim verschmelzenden Sinusende} \tag{43'}$$

sowohl für endliches wie unendliches p, wenn die Formeln im letzten Fall im limes $p \to \infty$ verstanden werden; für $y = \operatorname{arc\,tg} x$, $0 \leq x \leq \infty$ ist der Hauptwert $0 \leq y \leq \frac{\pi}{2}$ zu nehmen.

Bei der Abbildung eines h_a^0 in E_μ gilt $\lambda_\mu(\varrho) \equiv 1$; es bleibt also mit Definition 2 die Beziehung (43) gültig, wenn ein logarithmisches Elementargebiet an Stelle des randstelletragenden Sinusendes tritt.

Wird ein Flächenstück h_ε^0 durch $\zeta^{(\mu)}(w)$ in E_μ abgebildet, so können folgende zwei Fälle auftreten, wenn die Koordinaten der Windungspunkte in den beiden benachbarten Sinusenden der zugehörigen Knotenfolge überhaupt konvergieren:

a) beide Sinusenden sind randstelletragend,

b) eines ist verschmelzend, das andere randstelletragend [dann sind beide Konvergenzkoordinaten gleich].

Seien im Fall a) die Stärken der (relativ mittelbaren) Randstellen p' und p''; dann ist

$$\lambda_\mu(\varrho) = \frac{1}{\pi}\left\{\operatorname{arc\,tg}\frac{p'}{\pi} + \operatorname{arc\,tg}\frac{p''}{\pi}\right\} + O(\varrho^{-\alpha}). \tag{44}$$

Sei im Fall b) die Verschmelzungsstärke des einen Endes p', die Konvergenzstärke des anderen p'' [dann ist notwendig entweder $p' < p''$ oder, wenn $p' = p''$, das zum verschmelzenden Ende gehörige O-Glied kleiner als das zum randstelletragenden gehörige]. Man findet

$$\begin{cases} \lambda_\mu(\varrho) = \frac{1}{\pi}\left(\operatorname{arc\,tg}\frac{\pi}{p'} - \operatorname{arc\,tg}\frac{\pi}{p''}\right) + O(\varrho^{-\alpha}) \\ \quad = \frac{1}{\pi}\operatorname{arc\,tg}\frac{\pi(p''-p')}{\pi^2 + p'p''} + O(\varrho^{-\alpha}). \end{cases} \tag{44'}$$

Für $p' = p$, $p'' = \infty$ gehen (44), (44') in (43), (43') über. Der Fall, daß die Koordinaten der algebraischen Windungspunkte in einem oder in beiden Enden überhaupt nicht konvergieren, ist in den oben genannten Fällen a), b) als ausgearteter Fall mit $p' = 0$, $p'' = 0$ enthalten.

Die RIEMANNschen Flächen $W \in \mathfrak{W}$, deren Sinusenden nach den Definitionen 1, 3, 4 eine wohlbestimmte Stärke haben, fassen wir zur Klasse $\mathfrak{R}$ zusammen; nach einer Bemerkung auf S. 266 sind für sie die Voraussetzungen des Satzes 2a) erfüllt, wenn der (existierende) $\lim \sum \lambda_\mu(\varrho) > 0$ ist [denn für sie sind offenbar die Formeln (18) erfüllt]; ist hingegen $\lim \sum \lambda_\mu(\varrho) = 0$, so tritt Satz 3 ein.

Jedem logarithmischen Elementargebiet und jedem randstelletragenden Sinusende einer Fläche $W \in \mathfrak{R}$ ist genau eine Randstelle von W zugeordnet. Sind zwei derartige Randgebiete im Streckenkomplex durch ein verschmelzendes Sinusende getrennt, so ist beiden dieselbe Randstelle von W zugeordnet: in sie sind die zwei Relativ-Randstellen verschmolzen, die den beiden Randgebieten entsprechen. Die Anzahl der Randstellen von W ist somit gleich der Zahl der logarithmischen Elementargebiete und randstelletragenden Sinusenden vermindert um die Zahl der verschmelzenden Sinusenden, wenn das Ergebnis positiv ist, sonst gleich Eins oder Null, je nachdem ob es überhaupt ein logarithmisches Elementargebiet oder randstelletragendes Sinusende im Streckenkomplex gibt oder nicht[31]).

Wir betrachten eine Randstelle $\mathfrak{r}$ über a, sie sei entstanden durch Verschmelzen von $k \geq 1$ mittelbaren und unmittelbaren Relativ-Randstellen $\mathfrak{r}_\varkappa$ der

[31]) Für diesen Sachverhalt vgl. auch [7], insbesondere die Sätze 2, 3 dort, ebenso die folgende Fußnote 32.

Relativstärken $p_\varkappa$ ($\varkappa = 1, 2, \ldots, k$); sie seien der Reihe nach numeriert, wie die zugehörigen Randgebiete im Streckenkomplex aufeinander folgen. Zwei aufeinanderfolgende $\mathfrak{r}_\varkappa$, $\mathfrak{r}_{\varkappa+1}$ ($1 \leq \varkappa \leq k-1$) seien durch ein Sinusende der Verschmelzungsstärke $p_{\varkappa,\varkappa+1}$ verschmolzen — wie bereits bemerkt, ist $p_{\varkappa,\varkappa+1} \leq$ Min $(p_\varkappa, p_{\varkappa+1})$. Ehe wir nun die Gesamtstärke der Randstelle $\mathfrak{r}$ bestimmen, mögen zwei Fälle auseinandergehalten werden, der Fall der zerfällbaren und der der unzerfällbaren Flächen[32]). Ist die Fläche unzerfällbar, so ist die letzte Relativ-Randstelle $\mathfrak{r}_k$ mit der ersten $\mathfrak{r}_1$ durch ein Sinusende verschmolzen, es habe die Verschmelzungsstärke $p_{k,1}$; so gibt es k verschmelzende Sinusenden. Ist die Fläche zerfällbar, so sind die erste und die k-te Relativ-Randstelle nicht miteinander verschmolzen, es gibt alsdann nur $k-1$ verschmelzende Sinusenden in diesem Teil des Komplexes. Dieser letzte Fall tritt immer dann ein, wenn es wenigstens zwei verschiedene Randstellen oder einen logarithmischen Windungspunkt auf der Fläche W gibt. Der von der Randstelle $\mathfrak{r}$ herrührende Beitrag $\lambda^{\mathfrak{r}}(\varrho)$ zu $\sum \lambda_\mu(\varrho)$ ist dann gegeben durch $\sum_{\tau=1}^{t} \lambda^{\mathfrak{r}}_{\mu_\tau}(\varrho)$, wo $\pi \lambda^{\mathfrak{r}}_{\mu_\tau}(\varrho) \cdot \varrho$ die Länge desjenigen Teilbogens von $B^{\mu_\tau}_\varrho$ in der Viertelebene $\Im(\zeta_{\mu_\tau}) > 0$, $\Re(\zeta_{\mu_\tau}) \geq 0$ oder $\Re(\zeta_{\mu_\tau}) \leq 0$, für $f_{\mu_\tau} \in H$ ist, in deren Urbild auf W die Randstelle $\mathfrak{r}$ erreichbar ist. Danach ist

$$(45)\quad \begin{cases} \lambda^{\mathfrak{r}}(\varrho) = \lambda^{\mathfrak{r}}(p_1, p_k) + \dfrac{1}{\pi} \displaystyle\sum_{\varkappa=1}^{k-1} \operatorname{arc\,tg} \frac{\pi(p_\varkappa - p_{\varkappa,\varkappa+1})}{\pi^2 + p_\varkappa p_{\varkappa,\varkappa+1}} + \\ \qquad + \dfrac{1}{\pi} \displaystyle\sum_{\varkappa=2}^{k} \operatorname{arc\,tg} \frac{\pi(p_\varkappa - p_{\varkappa-1,\varkappa})}{\pi^2 + p_\varkappa p_{\varkappa-1,\varkappa}} + O(\varrho^{-\alpha}) = \lambda_{\mathfrak{r}} + O(\varrho^{-\alpha}), \end{cases}$$

dabei sei $\lambda_{\mathfrak{r}} = \lim \lambda^{\mathfrak{r}}(\varrho)$ und

$$(45')\quad \lambda^{\mathfrak{r}}(p_1, p_k) = \frac{1}{\pi}\left(\operatorname{arc\,tg} \frac{p_1}{\pi} + \operatorname{arc\,tg} \frac{p_\varkappa}{\pi}\right), \qquad \text{wenn } W \text{ zerfällbar,}$$

$$(45'')\quad \lambda^{\mathfrak{r}}(p_1, p_k) = \frac{1}{\pi}\left(\operatorname{arc\,tg} \frac{\pi(p_1 - p_{k,1})}{\pi^2 + p_1 p_{k,1}} + \operatorname{arc\,tg} \frac{\pi(p_k - p_{k,1})}{\pi^2 + p_k p_{k,1}}\right)$$

wenn W unzerfällbar[33]).

Für unendliche Zahlen $p_\varkappa$, $p_{\varkappa,\varkappa+1}$ ($\varkappa$ mod k) ist hier und im folgenden stets der Grenzwert der Ausdrücke zu verstehen, wenn $p_\varkappa$, $p_{\varkappa,\varkappa+1}$ gegen Unendlich

[32]) Eine Fläche heißt zerfällbar, wenn sie bei der Zerschneidung durch Ausstanzen über einer geeigneten geschlossenen doppelpunktfreien Kurve in wenigstens zwei unendlich vielblättrige Flächenstücke zerfällt. ULRICH [24] bemerkte die Bedeutung der Fallunterscheidung. Die Unzerfällbarkeit hat für eine Fläche $W \in \mathfrak{R}$ zur Folge, daß entweder keinem Randgebiet des Streckenkomplexes eine Randstelle entspricht, oder daß die Randgebiete des Streckenkomplexes von W, denen Randstellen entsprechen, alle dieselbe Koordinate haben, und daß alle anderen Randgebiete verschmelzende Sinusenden sind; insbesondere können Randstellen mit sich selbst verschmelzen (vgl. [5]).

[33]) Formal geht durch $p_{k,1} = 0$ die Formel (45′) aus (45″) hervor, also sozusagen durch Aufhören der Unzerfällbarkeit.

streben. Wir bemerken

$$\sum \lambda^{\mathfrak{r}}(\varrho) = \sum_{\mu=1}^{m} \lambda_\mu(\varrho) + O(\varrho^{-\alpha}), \tag{46}$$

wobei die erste Summe über alle Randstellen der Fläche W zu erstrecken ist und m die Anzahl aller Flächenstücke $f_\mu \in G \cup H$ [oder $f_{\mu_{\mathfrak{r}}} \in H$] ist, die bei der Zerschneidung der Fläche ausfallen: denn wenn die Länge des Stückes von B_ϱ^μ, das in der Viertelebene $\Im(\zeta_\mu) > 0$, $\Re(\zeta_\mu) \geq 0$ oder $\Re(\zeta_\mu) \leq 0$ nicht $O(\varrho^\alpha)$ ist, so ist in dem Urbild der Viertelebene gewiß eine Randstelle von W erreichbar.

Definition 5. Die durch die Formeln (45) gegebene Größe $\lambda_{\mathfrak{r}}$ ist die *Wachstumsstärke* der Randstelle $\mathfrak{r}$.

Die Wachstumsstärke $\lambda_{\mathfrak{r}}$ ist natürlich nicht negativ. Es gilt der

Satz 4[34]). a) *Jede nicht-mittelbare Randstelle — insbesondere also jede unmittelbare Randstelle — einer zerfällbaren Fläche $W \in \mathfrak{R}$ hat eine Wachstumsstärke $\lambda_{\mathfrak{r}} \geq 1$.*

b) *Jeder logarithmische Windungspunkt auf einer Fläche $W \in \mathfrak{R}$ hat die Wachstumsstärke Eins.*

c) *Auf (zerfällbaren oder unzerfällbaren) Flächen $W \in \mathfrak{R}$ gibt es mittelbare, auf unzerfällbaren Flächen $W \in \mathfrak{R}$ gibt es unmittelbare Randstellen der Wachstumsstärke Null.*

Beweis. a) Zu einer nicht-mittelbaren Randstelle einer zerfällbaren Fläche $W \in \mathfrak{R}$, die sich aus den benachbarten Relativ-Randstellen $\mathfrak{r}_1, \ldots, \mathfrak{r}_k$ der Relativstärken $p_1, \ldots, p_k$ $(k \geq 1)$ zusammensetzt, gehört wenigstens ein logarithmisches Elementargebiet im Streckenkomplex, also eine Relativ-Randstelle der Relativstärke ∞. Weil für die Verschmelzungsstärke $p_{\varkappa,\varkappa+1}$ des Sinusendes, das die Relativ-Randstellen $\mathfrak{r}_\varkappa, \mathfrak{r}_{\varkappa-1}$ $(1 \leq \varkappa \leq k-1)$ verschmilzt, die Beziehung $p_{\varkappa,\varkappa+1} \leq \operatorname{Min}(p_\varkappa, p_{\varkappa+1})$ gilt, folgt aus (45), (45′) die Ungleichung

$$\lambda_{\mathfrak{r}} \geq \frac{1}{\pi}\left\{\sum_{\varkappa=1}^{k-1} \operatorname{arc\,tg} \frac{|p_\varkappa - p_{\varkappa+1}|}{\pi^2 + p_\varkappa p_{\varkappa+1}} + \operatorname{arc\,tg} \frac{p_1}{\pi} + \operatorname{arc\,tg} \frac{p_k}{\pi}\right\}.$$

Weil wenigstens ein $p_\varkappa = \infty$ $(1 \leq \varkappa \leq k)$ ist, folgt die Behauptung.

b) Ein logarithmischer Windungspunkt besteht aus *einer* unmittelbaren Relativ-Randstelle, es ist $k = 1$ und $p_1 = \infty$; (45), (45′) geben wieder die Behauptung.

c) Eine mittelbare Randstelle, die aus einer einzigen Relativ-Randstelle der Relativstärke Null besteht, hat mit $k = 1$, $p_1 = 0$ nach (45), (45′) die Wachstums-

[34]) Ein Ergebnis MACLANES [11] über eine ähnliche Flächenklasse drückt sich in unserer Terminologie folgendermaßen aus: Für die Wachstumsstärke $\lambda_{\mathfrak{r}}$ einer mittelbaren Randstelle $\mathfrak{r}$ über der Koordinate ∞, die durch Konvergenz der Koordinaten a_n der algebraischen Windungspunkte in einem einfachen Sinusende entsteht, gilt $0 \leq \lambda_{\mathfrak{r}} \leq 1$; $\lambda_{\mathfrak{r}} = 1$ kann höchstens dann eintreten, wenn $\lim \frac{\log |a_n|}{n} = \infty$ ist. Die Formeln (45), (45′) ergeben in diesem Fall

$$\lambda_{\mathfrak{r}} = \frac{2}{\pi} \operatorname{arc\,tg} \frac{p}{\pi},$$ wenn p $(0 \leq p \leq \infty)$ die Konvergenzstärke des Sinusendes ist,

stärke Null, ebenso eine unmittelbare Randstelle auf einer unzerfällbaren Fläche für $k=1$, $p_1=p_{11}=\infty$, nach (45), (45″).

Für eine Fläche $W\in\mathfrak{R}$ erhält man sofort: entweder es ist $\lim\sum\lambda_\mu(\varrho)=0$, oder es gibt zwei positive Konstanten c_1, c_2, so daß $c_1\cdot\varrho<\nu(\varrho)<c_2\cdot\varrho+O(1)$. Mit (46) folgt aus Satz 2 [wenn $\lim\sum\lambda_\mu(\varrho)>0$] oder aus Satz 3 [wenn $\lim\sum\lambda_\mu(\varrho)=0$] also sofort

Satz 5. *Erzeugt $w(z)$ eine* RIEMANNsche *Fläche $W\in\mathfrak{R}$, die k unterschiedene Randstelle $\mathfrak{r}^{(\varkappa)}$ der Wachstumsstärken $\lambda_{\mathfrak{r}^{(\varkappa)}}$ ($\varkappa=1,2,\ldots,k$) trägt, so wächst $w(z)$ von der Ordnung $\frac{1}{2}\sum_{\varkappa=1}^{k}\lambda_{\mathfrak{r}^{(\varkappa)}}$, und zwar vom Mitteltypus dieser Ordnung, wenn $\sum_{\varkappa=1}^{k}\lambda_{\mathfrak{r}^{(\varkappa)}}>0$ ist. $\lambda_{\mathfrak{r}^{(\varkappa)}}$ bestimmt sich nach (45), (45′), (45″) aus dem Aufbau der Randstelle $\mathfrak{r}^{(\varkappa)}$.*

Korollar. *Zerfällbare* RIEMANNsche *Flächen $W\in\mathfrak{R}$ mit wenigstens einer nicht-mittelbaren Randstelle werden erst von Funktionen $w(z)$ einer Wachstumsordnung $s\geq\frac{1}{2}$ erzeugt.*

Das Korollar folgt sofort mit Satz 4a).

Satz 5 enthält offenbar den (allgemein bestmöglichen) Randstellensatz[35]) für die Flächenklasse $\mathfrak{R}$, der ja aussagt, daß eine Funktion $w(z)$, die die punktierte Ebene auf eine zerfällbare RIEMANNsche Fläche mit k unterschiedenen unmittelbaren Randstellen abbildet, wenigstens vom Mitteltyp der Ordnung $k/2$ anwächst. Für $W\in\mathfrak{R}$ ist die Aussage des Satzes 5 allerdings erheblich schärfer. Man erkennt ferner, weshalb es keinen dem Randstellensatz entsprechenden schärferen allgemeinen Satz geben kann, der außer unmittelbaren Randstellen auch mittelbare zuläßt: mittelbare Randstellen liefern nämlich unter Umständen gar keinen Beitrag zur Wachstumsordnung der erzeugenden Funktion[36]).

Als Beispiel betrachten wir die MITTAG-LEFFLERschen E_α-Funktionen $E_\alpha(z)=\sum_{\nu=0}^{\infty}\frac{z^\nu}{\Gamma(1+\alpha\nu)}$ für $\alpha\geq 2$. Die erzeugten RIEMANNschen Flächen[37]) haben eine einzige unmittelbare Randstelle, und zwar über ∞, die Verzweigungspunkte liegen sämtlich über der reellen Achse; der Streckenkomplex S^* wird gebildet von einem Sinusende,

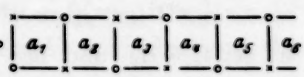


Fig. 5

das in ein logarithmisches Elementargebiet hineinragt, der Streckenkomplex hat also zwei Randgebiete: ein Sinusgebiet und ein logarithmisches Elementargebiet (Fig. 5). Für die Koordinaten a_n der algebraischen Windungspunkte

[35]) Siehe z. B. ULLRICH [24].

[36]) Eine feinere Unterscheidung in der Definition der Konvergenzstärke und der Verschmelzungsstärke von Sinusenden verbunden mit genaueren Abschätzungen in § 6 würde auch den Einfluß von Randstellen auf den Typus einer Wachstumsordnung erkennen lassen; vgl. das Beispiel der Gammafunktion in [6].

[37]) Vgl. die Untersuchung von FISCHER [3].

des Sinusendes gilt[38])

$$a_n = (-1)^n \cdot \frac{2}{\alpha} \cdot \sin\frac{\pi}{\alpha}\, e^{\left(n+\frac{1}{2}-\frac{1}{\alpha}\right)\pi \operatorname{ctg}\frac{\pi}{\alpha}} \left(1+\varepsilon(n)\right);$$

das Sinusende hat also die Verschmelzungsstärke $\pi \operatorname{ctg}\frac{\pi}{\alpha}$. Die Fläche gehört zur Klasse $\mathfrak{R}$, sie ist unzerfällbar für $\alpha > 2$, daher ist die Wachstumsstärke λ_∞ der Randstelle über ∞ nach (45), (45″) gegeben durch $2 \cdot \frac{1}{\pi} \operatorname{arc\,tg} \frac{\pi(p_1 - p_{11})}{\pi^2 + p_1 p_{11}}$. Mit $p_1 = \infty$, $p_{11} = \pi \operatorname{ctg}\frac{\pi}{\alpha}$ erhält man

$$\lambda_\infty = 2 \cdot \frac{1}{\pi} \operatorname{arc\,tg} \frac{1}{\operatorname{ctg}\frac{\pi}{\alpha}} = \frac{2}{\alpha}. \tag{47}$$

Für $\alpha = 2$ wird die unmittelbare Randstelle zum logarithmischen Windungspunkt und die Fläche wird demgemäß zerfällbar [es ist $E_2(z) = \cos h\sqrt{z}$], also ist nach Satz 4b) $\lambda_\infty = 1$; es gilt also (47) auch für $\alpha = 2$. Nach Satz 5 folgt das bekannte Ergebnis: die Wachstumsordnung von $E_\alpha(z)$ [für $\alpha \geqq 2$] ist α^{-1}.

§ 8. Defekt der Randstellensorten

Der Defekt $\delta(a)$ eines Wertes a bestimmt sich für die Funktionen $w(z)$, die Flächen $W \in \mathfrak{R}$ erzeugen, aus den Anzahlen $\nu(\varrho, a)$ und $\nu(\varrho)$ nach Satz 2 [Teil b) insbesondere], *wenn die Wachstumsordnung von $w(z)$ positiv* ist[39]); denn wie sich gleich ergeben wird, ist für große ϱ entweder $c_a \cdot \varrho = \mathsf{N}(\varrho, a) + O(\varrho^{1-\alpha})$ mit $c_a > 0$ — dann sind die ersten der Gln. (36), (36′) erfüllt — oder es ist $\nu(\varrho, a) = o(\varrho)$; dann ermöglicht die Formel (38) eine exakte Aussage: es ist nämlich jedenfalls $c \cdot \varrho < \mathsf{N}(\varrho) + O(1)$ mit positivem c. Es sei also $W \in \mathfrak{R}$.

Wir bestimmen zunächst den von einer Ebene E_{μ_τ} herrührenden Beitrag $\nu^{(\mu_\tau)}(\varrho)$ bzw. $\nu^{(\mu_\tau)}(\varrho, a)$ zu $\nu(\varrho)$ bzw. $\nu(\varrho, a)$. Es sei Q_{μ_τ} in E_{μ_τ} Bild eines Flächenstückes $f^0_{\mu_\tau}$ mit $f_{\mu_\tau} \in H$, und es seien $w'_{\mu_\tau}, w''_{\mu_\tau}$ die Koordinaten des linken und des rechten Randgebietes der zugehörigen Knotenfolge. Ist $f_{\mu_\tau} \in H_a$, so ist

$$\text{(48a)} \quad \begin{cases} \nu^{(\mu_\tau)}(\varrho) = \dfrac{\varrho}{2\pi} + O(1), \quad \nu^{(\mu_\tau)}(\varrho, a) = \dfrac{\varrho}{2\pi} + O(1) \quad \text{für} \quad a \neq w'_{\mu_\tau}, w''_{\mu_\tau}; \\ \nu^{(\mu_\tau)}(\varrho, w'_{\mu_\tau}) = \nu^{(\mu_\tau)}(\varrho, w''_{\mu_\tau}) = 0. \end{cases}$$

Es sei $f_{\mu_\tau} \in H_b$, und das Sinusgebiet liege im linken Teil von f_{μ_τ}. Dann ist

$$\text{(48b)} \quad \begin{cases} \nu^{(\mu_\tau)}(\varrho) = \dfrac{\varrho}{2\pi} \cdot \sin \pi \lambda_{\mu_\tau}(\varrho) + O(1) & \text{für} \quad \lambda_{\mu_\tau}(\varrho) \leqq \dfrac{1}{2}, \\ \quad = \dfrac{\varrho}{2\pi} + O(1) & \text{für} \quad \lambda_{\mu_\tau}(\varrho) \geqq \dfrac{1}{2}; \\ \nu^{(\mu_\tau)}(\varrho, a) = \nu^{(\mu_\tau)}(\varrho) + O(1) & \text{für} \quad a \neq w'_{\mu_\tau}, w''_{\mu_\tau}, \\ \nu^{(\mu_\tau)}(\varrho, w'_{\mu_\tau}) = \dfrac{\varrho}{2\pi} \cdot \sin \pi \lambda_{\mu_\tau}(\varrho) + O(1), \quad \nu^{(\mu_\tau)}(\varrho, w''_{\mu_\tau}) = 0. \end{cases}$$

[38]) FISCHER [3], Teil II, Formel (33).

[39]) Auch im Fall $\lim \sum \lambda_\mu(\varrho) = 0$ würden sich bei genauerem Abschätzen in § 6 Ergebnisse erzielen lassen; z. B. würde man zeigen können, daß es Funktionen der Ordnung Null gibt, für die endlich aber beliebig viele Stellen positiven Defekt (der Summe 2) haben.

Ganz analoge Formeln erhält man, wenn das Sinusgebiet im rechten Teil von f_{μ_τ} liegt: es ist dann $\lambda_{\mu_\tau}(\varrho)$ durch $1-\lambda_{\mu_\tau}(\varrho)$ zu ersetzen und w'_{μ_τ} mit w''_{μ_τ} zu vertauschen. Ist jedoch $f_{\mu_\tau}\in H_c$, so verbinden wir aufeinanderfolgende Grätenendpunkte durch ein Geradenstück. Die Gleichung des gewonnenen Streckenzuges sei in Polarkoordinaten

$$\varphi_{\mu_\tau}=\pi\,\lambda^{(r)}_{\mu_\tau}(\varrho)\quad\text{bzw.}\quad\varphi_{\mu_\tau}=\pi\,\lambda^{(l)}_{\mu_\tau}(\varrho)\quad[\lambda^{(r)}_{\mu_\tau}(\varrho)<\lambda^{(l)}_{\mu_\tau}(\varrho)]$$

für die Grätenendpunkte auf dem rechten bzw. dem linken Rand von Q_{μ_τ}. Es ist dann

$$\text{(48c)}\quad\begin{cases}\nu^{(\mu_\tau)}(\varrho)=\dfrac{\varrho}{2\pi}+O(1),\quad\text{wenn}\quad\lambda^{(r)}_{\mu_\tau}(\varrho)\leq\dfrac{1}{2}\quad\text{und}\quad\lambda^{(l)}_{\mu_\tau}(\varrho)\geq\dfrac{1}{2}\\ \qquad=\dfrac{\varrho}{2\pi}\,\mathrm{Max}\left\{\sin\pi\,\lambda^{(r)}_{\mu_\tau}(\varrho),\ \sin\pi\,\lambda^{(l)}_{\mu_\tau}(\varrho)\right\}+O(1)\quad\text{sonst};\\ \nu^{(\mu_\tau)}(\varrho,a)=\nu^{(\mu_\tau)}(\varrho)+O(1)\quad\text{für}\quad a\neq w'_{\mu_\tau},\,w''_{\mu_\tau},\\ \nu^{(\mu_\tau)}(\varrho,w'_{\mu_\tau})=\dfrac{\varrho}{2\pi}\sin\pi\,\lambda^{(l)}_{\mu_\tau}(\varrho)+O(1),\\ \nu^{(\mu_\tau)}(\varrho,w''_{\mu_\tau})=\dfrac{\varrho}{2\pi}\sin\pi\,\lambda^{(r)}_{\mu_\tau}(\varrho)+O(1).\end{cases}$$

Wir spalten die Untersuchung wieder auf in den Fall der zerfällbaren und den der unzerfällbaren Flächen. Bei den letzteren kann als einziger Wert die Randstellenkoordinate defekt sein, denn die Anzahlfunktion $\nu(\varrho,a)$ unterscheidet sich für alle anderen Werte von a nur um $O(1)$ von $\nu(\varrho)$, wie man mit Hilfe von (42), (42'), (43'), (48b), (48c) erschließt[40]). Wir wollen, wie eingangs erwähnt, annehmen, es sei $\lim\sum_{\mu=1}^{m}\lambda_\mu(\varrho)=\Lambda>0$, d.i.: die Wachstumsordnung der erzeugenden Funktion sei positiv.

Sei $W\in\mathfrak{R}$ eine unzerfällbare Fläche mit einer Randstelle $\mathfrak{r}$ über w_0, die sich aus n mittel- und unmittelbaren Relativ-Randstellen $\mathfrak{r}_\nu$ mit den Relativstärken p_ν $(\nu=1,2,\ldots,n)$ zusammensetzt; die benachbarten Relativ-Randstellen $\mathfrak{r}_\nu$, $\mathfrak{r}_{\nu+1}$ seien verschmolzen durch ein Sinusende der Verschmelzungsstärke $p_{\nu,\nu+1}$ $(\nu \bmod n)$. Es ist $\mathrm{Min}(p_\nu,p_{\nu+1},p_{\nu,\nu+1})=p_{\nu,\nu+1}$. Man entnimmt aus der Definition von $\mathsf{N}(\varrho)$ den Formeln (48) nach den Definitionen der Relativstärke von Relativ-Randstellen und der Verschmelzungsstärke von Sinusenden die Beziehungen

$$\text{(49)}\qquad\mathsf{N}(\varrho)=\frac{2\varrho}{\Lambda}\sum_{\nu=1}^{n}\frac{1}{\sqrt{p^2_{\nu,\nu+1}+\pi^2}}+O(\varrho^{1-\alpha}),$$

$$\text{(49')}\quad\begin{cases}\nu(\varrho,w_0)=\varrho\cdot\displaystyle\sum_{\nu=1}^{n}\frac{1}{\sqrt{p^2_\nu+\pi^2}}+O(\varrho^{1-\alpha}),\\ \mathsf{N}(\varrho,w_0)=\dfrac{2\varrho}{\Lambda}\cdot\displaystyle\sum_{\nu=1}^{n}\frac{1}{\sqrt{p^2_\nu+\pi^2}}+O(\varrho^{1-\alpha})\end{cases}$$

[40]) Vgl. Fußnote 32.

[wegen $\lim \sum \lambda_\mu(\varrho) = \Lambda > 0$ ist $p_{\nu,\nu+1} \neq \infty$ für wenigstens einen Wert ν]. Der Defekt $\delta(w_0)$ der Stelle w_0 ist, nach (37) für $\sum p_\nu^{-1} > 0$, nach (38) für $\sum p_\nu^{-1} = 0$, also gegeben durch

$$\delta(w_0) = \frac{\sum_{\nu=1}^{n} (p_{\nu,\nu+1}^2 + \pi^2)^{-\frac{1}{2}} - \sum_{\nu=1}^{n} (p_\nu^2 + \pi^2)^{-\frac{1}{2}}}{\sum_{\nu=1}^{n} (p_{\nu,\nu+1}^2 + \pi^2)^{-\frac{1}{2}}}. \tag{49''}$$

Für eine unmittelbare Relativ-Randstelle $\mathfrak{r}_\nu$ gilt $p_\nu = \infty$, für eine mittelbare gilt (ν mod n)

$$p_\nu \geq \operatorname{Max}(p_{\nu-1,\nu}, p_{\nu,\nu+1}), \tag{50}$$

dies ist aber auch die einzige Einschränkung. Sind alle $\mathfrak{r}_\nu$ relativ unmittelbar [$p_\nu = \infty$], so ist $\delta(w_0) = 1$; dagegen kann $\delta(w_0)$ der Null beliebig nahe kommen, wenn neben einer festen Anzahl von unmittelbaren Relativ-Randstellen eine genügend große Zahl von sehr schwachen mittelbaren Relativ-Randstellen — etwa solchen, für die in (50) das Gleichheitszeichen steht — die Randstelle $\mathfrak{r}$ bildet. Offenbar kann $\delta(w_0)$ durch (mannigfache) Wahl der Stärken der Relativ-Randstellen gleich einer vorgegebenen Zahl d, $0 < d \leq 1$, gemacht werden. Es gilt also der

Satz 6. *Zu jeder vorgegebenen Zahl d, $0 < d \leq 1$, gibt es eine (unzerfällbare) Fläche $W \in \mathfrak{R}$, deren erzeugende Funktion $w(z)$ als einzigen defekten Wert den Wert w_0 mit dem Defekt $\delta(w_0) = d$ hat.*

Sei $W \in \mathfrak{R}$ nun zerfällbar. Es werde eine Randstelle $\mathfrak{r}$ über a gebildet durch n Relativ-Randstellen $\mathfrak{r}_\nu$ der Stärken p_ν ($\nu = 1, 2, \ldots, n$), die verschmolzen seien durch $n-1$ Sinusenden der Verschmelzungsstärken $p_{\nu,\nu+1}$ ($\nu = 1, 2, \ldots, n-1$). Mit $\mathsf{N}_\mathfrak{r}(\varrho)$ bzw. $\mathsf{N}_\mathfrak{r}(\varrho, a)$ bezeichnen wir den Beitrag zu $\mathsf{N}(\varrho)$ bzw. $\mathsf{N}(\varrho, a)$, der von denjenigen Flächenstücken $f_\mu \in H$ herrührt, die eine der Relativ-Randstellen $\mathfrak{r}_\nu$ am Rand haben. Es ist

$$\mathsf{N}_\mathfrak{r}(\varrho) = \frac{2\varrho}{\Lambda} \left\{ \sum_{\nu=1}^{n-1} (p_{\nu,\nu+1}^2 + \pi^2)^{-\frac{1}{2}} + \frac{1}{\pi} \right\} + O(\varrho^{1-\alpha}). \tag{51}$$

Formal kann man (51) mit (49) dadurch in Beziehung bringen, daß man den letzten Summanden $1/\pi$ in der Klammer auffaßt als herrührend von dem fehlenden n-ten verschmelzenden Ende, was sich auswirkt wie ein Ende der Verschmelzungsstärke Null. Weiter gilt

$$\mathsf{N}_\mathfrak{r}(\varrho, a) = \frac{2\varrho}{\Lambda} \sum_{\nu=1}^{n} (p_\nu^2 + \pi^2)^{-\frac{1}{2}} + O(\varrho^{1-\alpha}). \tag{51'}$$

Es ist nun zu beachten, daß die einfach über alle Randstellen erstreckte Summe der $\mathsf{N}_\mathfrak{r}(\varrho)$ die zu einer Ebene E_μ gehörige Anzahlfunktion $\int \nu^{(\mu)}(\varrho)\, d\mathfrak{f}(\varrho)$ doppelt enthält, wenn im zugehörigen Flächenstück $f_\mu \in H$ zwei verschiedene Randstellen von W erreichbar sind; es ist hierbei

$$\int \nu^{(\mu)}(\varrho)\, d\mathfrak{f}(\varrho) = \frac{2}{\Lambda} \frac{\varrho}{2\pi} + O(\varrho^{1-\alpha}).$$

Umgekehrt sind die Sinusenden, in denen keine Konvergenz stattfindet, unberücksichtigt geblieben. Wir treffen daher folgende Verabredung: Ein Sinusgebiet, in dem die Koordinaten der Elementargebiete nicht konvergieren, sei aufgefaßt als ob es eine Randstelle bewirke, der aber keine Koordinate zugeordnet sei. Dann ist

$$\mathsf{N}(\varrho) = \sum_{\mathfrak{r}} \left(\mathsf{N}_{\mathfrak{r}}(\varrho) - \frac{2}{\Lambda} \cdot \frac{\varrho}{2\pi} \right) + O(\varrho^{1-\alpha}), \tag{52}$$

wobei die Summation unter Beachtung der Verabredung über alle Randstellen zu erstrecken ist. Ferner ist

$$\mathsf{N}(\varrho, a) = \sum_{r=a} \left(\mathsf{N}_{\mathfrak{r}}(\varrho, a) - \frac{2}{\Lambda} \frac{\varrho}{2\pi} \right) + \sum_{r \neq a} \left(\mathsf{N}_{\mathfrak{r}}(\varrho) - \frac{2}{\Lambda} \frac{\varrho}{2\pi} \right) + O(\varrho^{1-\alpha}), \tag{52'}$$

wobei die erste Summe über alle Randstellen $\mathfrak{r}$ der Koordinate $r = a$, die zweite über alle übrigen zu erstrecken ist. Denn $\sum_{r=a} \mathsf{N}_{\mathfrak{r}}(\varrho, a) + \sum_{r \neq a} \mathsf{N}_{\mathfrak{r}}(\varrho)$ zählt in allen Halbebenen, die zu *zwei* Randstellen $\mathfrak{a}_1, \mathfrak{a}_2$ gehören, von denen keine über a liegt, die a-Stellen doppelt [vgl. die Bestimmung von $\mathsf{N}(\varrho)$ oben]; daher ist in jedem Summanden der zweiten Summe die Korrektur $-\frac{2}{\Lambda}\frac{\varrho}{2\pi}$ anzubringen; in allen Halbebenen, die zu zwei Randstellen gehören, von denen eine über a liegt, zählt $\sum_{r=a} \mathsf{N}_{\mathfrak{r}}(\varrho, a)$ die richtige Anzahl, $\sum_{r \neq a} \mathsf{N}_{\mathfrak{r}}(\varrho)$ zählt hingegen bis auf $O(\varrho^{1-\alpha})$ zusätzlich $\frac{2}{\Lambda}\frac{\varrho}{2\pi}$; diesen Überschuß berücksichtigt das Korrekturglied $-\frac{2}{\Lambda}\frac{\varrho}{2\pi}$ im ersten Summanden. Rechnet man $\mathsf{N}(\varrho)$, $\mathsf{N}(\varrho, a)$ nach (51), (51'), (52), (52') aus, so erhält man ähnlich wie im Fall unzerfällbarer Flächen nach (37) bzw. (38) in Satz 2 — der anwendbar ist wegen der Annahme $\lim \sum \lambda_\mu(\varrho) = \Lambda > 0$ — den

Satz 7. *Die meromorphe Funktion $w(z)$ positiver Wachstumsordnung erzeuge eine zerfällbare* RIEMANNsche *Fläche $W \in \mathfrak{R}$ mit k verschiedenen Randstellen $\mathfrak{r}^{(\varkappa)}$ ($\varkappa = 1, 2, \ldots, k$), wobei jedes Sinusende, in dem die Koordinaten der Windungspunkte nicht konvergieren, als gesonderte Randstelle ohne Koordinate aufgefaßt sei. Die Randstelle $\mathfrak{r}^{(\varkappa)}$ sei gebildet von $n_\varkappa$ Relativ-Randstellen $\mathfrak{r}_\nu^{(\varkappa)}$ ($\nu = 1, 2, \ldots, n_\varkappa$) der Relativstärken $p_{\varkappa_\nu}$, die durch $n_\varkappa - 1$ Sinusenden der Verschmelzungsstärken $p_{\varkappa_{\nu,\nu+1}}$ ($\nu = 1, 2, \ldots, n_\varkappa - 1$) verschmolzen sind. Bezeichnen wir die Spur von $\mathfrak{r}^{(\varkappa)}$ mit $r_\varkappa$, so ist der Defekt $\delta(a)$ der Stellensorte a gegeben durch*

$$\begin{cases} \delta(a) = 1 - \dfrac{\displaystyle\sum_{\substack{\varkappa \\ r_\varkappa \neq a}} \left\{ \sum_{\nu=1}^{n_\varkappa - 1} (p_{\varkappa_{\nu,\nu+1}}^2 + \pi^2)^{-\frac{1}{2}} + \frac{1}{2\pi} \right\} + \sum_{\substack{\varkappa \\ r_\varkappa = a}} \left\{ \sum_{\nu=1}^{n_\varkappa} (p_{\varkappa_\nu}^2 + \pi^2)^{-\frac{1}{2}} - \frac{1}{2\pi} \right\}}{\displaystyle\sum_{\varkappa=1}^{k} \left\{ \sum_{\nu=1}^{n_\varkappa - 1} (p_{\varkappa_{\nu,\nu+1}}^2 + \pi^2)^{-\frac{1}{2}} + \frac{1}{2\pi} \right\}} \\ = \dfrac{\displaystyle\sum_{\substack{\varkappa \\ r_\varkappa = a}} \left\{ \sum_{\nu=1}^{n_\varkappa - 1} (p_{\varkappa_{\nu,\nu+1}}^2 + \pi^2)^{-\frac{1}{2}} + \frac{1}{\pi} - \sum_{\nu=1}^{n_\varkappa} (p_{\varkappa_\nu}^2 + \pi^2)^{-\frac{1}{2}} \right\}}{\displaystyle\sum_{\varkappa=1}^{k} \left\{ \sum_{\nu=1}^{n_\varkappa - 1} (p_{\varkappa_{\nu,\nu+1}}^2 + \pi^2)^{-\frac{1}{2}} + \frac{1}{2\pi} \right\}} \end{cases} \tag{53}$$

Es sei $\mathfrak{r}^{(\varkappa)}$ eine Randstelle, die gebildet wird von $n_\varkappa$ Relativ-Randstellen $\mathfrak{r}_\nu^{(\varkappa)}$ $(\nu = 1, 2, \ldots, n_\varkappa)$ der Relativstärken $p_{\varkappa_\nu}$, wobei $\mathfrak{r}_\nu^{(\varkappa)}$, $\mathfrak{r}_{\nu+1}^{(\varkappa)}$ verschmolzen seien durch ein Sinusende der Verschmelzungsstärke $p_{\varkappa_{\nu,\nu+1}}$ (ν mod $n_\varkappa$) [formal werde bei zerfällbaren Flächen für $\nu = n_\varkappa$ die Größe $p_{\varkappa_{\nu,\nu+1}} = 0$ gesetzt]. Die Formeln (45) zeigen, daß $\mathfrak{r}^{(\varkappa)}$ dann und nur dann einen Beitrag zur Wachstumsordnung gibt, wenn nicht alle Differenzen $p_{\varkappa_\nu} - p_{\varkappa_{\nu,\nu+1}}$ und $p_{\varkappa_\nu} - p_{\varkappa_{\nu-1,\nu}}$ gleich Null sind; die Formeln (49''), (53) zeigen, daß dieselbe Bedingung für den Beitrag zum Defekt gilt [in den Differenzen ist beidemale formal $\infty - \infty = 0$ zu setzen], *wenn auch noch* $\sum\sum p_{\varkappa_{\nu,\nu+1}}^{-1} > 0$ ist. Berücksichtigen wir, daß die letzte Bedingung gleichbedeutend damit ist, daß die zugehörige meromorphe Funktion positive Wachstumsordnung hat, so folgt der

Satz 8. *Ist auf einer Fläche $W \in \mathfrak{R}$ eine Randstelle so kräftig, daß sie positive Wachstumsstärke hat, so hat die Stellensorte der Randstelle positiven Defekt. Umgekehrt: Gibt eine Randstelle einen positiven Beitrag zum Defekt ihrer Stellensorte, so ist auch ihre Wachstumsstärke positiv, wenn die Wachstumsordnung der erzeugenden meromorphen Funktion nicht überhaupt Null ist.*

Der erste Teil des Satzes zeigt, daß auch eine mittelbare Randstelle einen positiven Beitrag zum Defekt ihrer Stellensorte bewirken kann; dies ist in der Regel so, nämlich wenn die mittelbare Randstelle nicht die Wachstumsstärke Null hat[41]). Die Einschränkung im zweiten Teil des Satzes ist notwendig, denn selbstverständlich gibt es ganze Funktionen der Ordnung Null, bei denen also $\delta(\infty) = 1$ [vgl. auch Fußnote 39].

Für die Summe aller Defekte erhalten wir durch Summation über a aus (53) die Beziehung

$$\sum \delta(a) = 1 + \frac{\frac{k}{2\pi} - \sum_{\varkappa=1}^{k} \sum_{\nu=1}^{n_\varkappa} (p_{\varkappa_\nu}^2 + \pi^2)^{-\frac{1}{2}}}{\frac{k}{2\pi} + \sum_{\varkappa=1}^{k} \left\{ \sum_{\nu=1}^{n_\varkappa - 1} (p_{\varkappa_{\nu,\nu+1}}^2 + \pi^2)^{-\frac{1}{2}} \right\}}. \tag{54}$$

Man sieht, die Defektsumme kann dann und nur dann gleich 2 werden, wenn in (54) in allen Summanden der Doppelsumme sowohl alle $p_{\varkappa_{\nu,\nu+1}}$ wie alle $p_{\varkappa_\nu}$ unendlich sind. [Dann ist übrigens auch jeder Defekt eine rationale Zahl.] Der Grund liegt darin, daß ein von unendlich verschiedenes $p_{\varkappa_{\nu,\nu+1}}$ oder $p_{\varkappa_\nu}$ in den zugehörigen Ebenen E_μ Grätenendpunkte in solcher asymptotischer Anzahl mit sich bringt, daß die von den zugehörigen algebraischen Windungspunkten herrührende gesamte algebraische Verzweigtheit[42]) positiv ausfällt.

Es gelingt nun, Funktionen zu konstruieren, die an k vorgegebenen Stellen $a_\varkappa$ $(\varkappa = 1, 2, \ldots, k)$ der reellen Achse den vorgegebenen Defekt $\delta(a_\varkappa) = d_\varkappa$ besitzen, wenn die Bedingungen $d_\varkappa < 1$, $\sum_{\varkappa=1}^{k} d_\varkappa < 2$ erfüllt sind. Zur Konstruktion — ich will nicht auf Einzelheiten eingehen, ebensowenig darauf, daß die Realität der $a_\varkappa$ natürlich unwesentlich ist — wählen wir zunächst positive Zahlen

[41]) Vgl. Fußnote 5.

[42]) Dies ist der Gesamtindex Φ in ULLRICH [21].

$p_{i_{\nu,\nu+1}}$[43]) ($\nu = 1, 2, \ldots, n_i - 1$), so daß die aus diesen Zahlen [vgl. (53)] bestimmten Größen $d'_\varkappa$, $d''_\varkappa$ mit

$$d'_\varkappa = \frac{\sum\limits_{\substack{i \\ r_i = a_\varkappa}} \frac{1}{\pi}}{\sum\limits_{i=1}^{i} \left\{ \sum\limits_{\nu=1}^{n_i-1} (p^2_{i_{\nu,\nu+1}} + \pi^2)^{-\frac{1}{2}} + \frac{1}{2\pi} \right\}}, \quad d''_\varkappa = \frac{\sum\limits_{\substack{i \\ r_i = a_\varkappa}} \left\{ \frac{1}{\pi} + \sum\limits_{\nu=1}^{n_i-1} (p^2_{i_{\nu,\nu+1}} + \pi^2)^{-\frac{1}{2}} \right\}}{\sum\limits_{i=1}^{i} \left\{ \sum\limits_{\nu=1}^{n_i-1} (p^2_{i_{\nu,\nu+1}} + \pi^2)^{-\frac{1}{2}} + \frac{1}{2\pi} \right\}}$$

die Beziehungen $d'_\varkappa < d_\varkappa < d''_\varkappa$, $\sum\limits_{\varkappa=1}^{k} d''_\varkappa < 2$ erfüllen; eine solche Wahl ist offenbar möglich. Jetzt können dazu noch Zahlen p_{i_ν} ($\nu = 1, 2, \ldots, n_i$) so bestimmt werden[43]), daß gilt

$$\frac{\sum\limits_{\substack{i \\ r_i = a_\varkappa}} \left\{ \sum\limits_{\nu=1}^{n_i-1} (p^2_{i_{\nu,\nu+1}} + \pi^2)^{-\frac{1}{2}} + \frac{1}{\pi} - \sum\limits_{\nu=1}^{n_i} (p^2_{i_\nu} + \pi^2)^{-\frac{1}{2}} \right\}}{\sum\limits_{\varkappa=1}^{k} \left\{ \sum\limits_{\nu=1}^{n_i-1} (p^2_{i_{\nu,\nu+1}} + \pi^2)^{-\frac{1}{2}} + \frac{1}{2\pi} \right\}} = d_\varkappa .$$

Die Funktion, die die Fläche mit den Randstellen $\mathfrak{r}_i$ über der Koordinate r_i der zugehörigen Relativstärken p_{i_ν} und Verschmelzungsstärken $p_{i_{\nu,\nu+1}}$ erzeugt, löst dann die Aufgabe. Es gilt also der

Satz 9. *Es sei eine Defektverteilung* $\delta(a_\varkappa) = d_\varkappa$ ($\varkappa = 1, 2, \ldots, k$) *vorgegeben mit* $0 < d_\varkappa < 1$, $\sum\limits_{\varkappa=1}^{k} d_\varkappa < 2$ *und* $a_\varkappa$ *reell. Es gibt dann eine (eine Fläche* $W \in \mathfrak{R}$ *erzeugende) meromorphe Funktion* $w(z)$, *die die vorgegebene Defektverteilung besitzt.*

Die Ordnung einer Funktion $w(z)$, die eine Fläche $W \in \mathfrak{R}$ erzeugt, kann nicht unmittelbar aus der Defektverteilung gewonnen werden; denn in die Ordnung von $w(z)$ gehen die Summe von Arcustangenten gewisser Ausdrücke ein, die aus den Relativstärken und Verschmelzungsstärken gebildet sind, während die Defekte algebraisch von den Relativ- und den Verschmelzungsstärken abhängen; und es ist nicht möglich, aus der Größe dieser direkt auf die Größe jener, d.i. auf die Ordnung, zu schließen. Dennoch besteht natürlich ein gewisser Zusammenhang. Über bekannte Aussagen hinausgehend vermutete TEICHMÜLLER [*20*], daß für jede meromorphe eine zerfällbare Fläche erzeugende Funktion $w(z)$ mit den Defekten $\delta(a_\varkappa) = d_\varkappa$ ($\varkappa = 1, 2, \ldots, k$) die Relation

$$s \geq \frac{1}{\pi} \sum_{\varkappa=1}^{k} \arccos(1 - d_\varkappa)$$

gelte, wo s die (obere) Ordnung von $w(z)$ ist. Als „Testobjekt" für diese Vermutung würden sich die Flächen $W \in \mathfrak{R}$ etwa eignen.

[43]) $p_{i_{\nu,\nu+1}}$, p_{i_ν} gehöre zur Randstelle $\mathfrak{r}^{(i)}$ ($i = 1, 2, \ldots, i$) des Spurpunktes r_i. Die Randstellen werden in geeigneter Weise in die Stellensorten $a_\varkappa$ gelegt.

§ 9. Die Abschätzungsmethode von Ahlfors

Ist die Fläche W zu unregelmäßig gebaut, als daß (16) durch (16') fortgesetzt werden könnte, so greifen wir, um ein Ergebnis zu erzielen, zurück auf ein Verfahren, das sich auf Überlegungen gründet, die zuerst von AHLFORS [1] angestellt wurden. Die Resultate, die wir auf diesem Wege erhalten, sind im allgemeinen ungenauer, jedoch erlaubt die AHLFORSsche Methode, den Satz 3 zu beweisen und so die Lücke zu schließen, die Satz 2 für Funktionen der Ordnung Null offen ließ. Wir übernehmen die Bezeichnungen der vorangehenden Paragraphen.

Erzeuge also die meromorphe Funktion $w(z)$ die Fläche $W \in \mathfrak{W}$ als Bild der endlichen z-Ebene, $B_{\varrho,z}$ sei in der z-Ebene das Urbild der geschlossenen Kurve $B_{\varrho,w}$ auf W, und es sei ϱ hinreichend groß. Wir bilden die längs der positiv reellen Achse aufgeschnittene z-Ebene ab durch $v = \log z$ auf den Streifen $0 \leq \Im(v) \leq 2\pi$. Weiter setzen wir

$$\omega(\varrho) = \log \frac{u_2(\varrho)}{u_1(\varrho)}.$$

Die v-Bildkurve $B_{\varrho,v}$ von $B_{\varrho,z}$ hat eine Länge, die sicher nicht kleiner ist als die Diagonale des Rechtecks mit den Seiten $\omega(\varrho)$ und 2π:

$$\sqrt{4\pi^2 + \omega^2(\varrho)} \leq \sum_{\mu=1}^{m} \int_{B_\varrho^\mu} \left| \frac{dv}{d\zeta_\mu} \right| |d\zeta_\mu|,$$

wobei m wieder die Gesamtzahl der bei der Zerschneidung von W ausfallenden Flächenstücke $f_\mu \in G \cup H$ ist. Nach der SCHWARZschen Ungleichung folgt

$$(55) \quad 4\pi^2 + \omega^2(\varrho) \leq \left(\sum_{\mu=1}^{m} \int_{B_\varrho^\mu} D_\mu(\zeta_\mu) |d\zeta_\mu| \right) \left(\sum_{\mu=1}^{m} \int_{B_\varrho^\mu} \left| \frac{dv}{d\zeta_\mu} \right|^2 \cdot \frac{1}{D_\mu(\zeta_\mu)} \cdot |d\zeta_\mu| \right),$$

wo $D_\mu(\zeta_\mu)$ der Dilatationsquotient der quasikonformen Abbildung $q_\mu(\zeta_\mu)$ ist. Nun ist

$$\int_{B^\mu} D_\mu |d\zeta_\mu| = \pi \lambda_\mu(\varrho) \cdot \varrho + \pi \varkappa_\mu(\varrho) \quad \text{mit nicht-negativem } \varkappa_\mu(\varrho) = O(1)$$

sobald ϱ genügend groß; wir erhalten somit nach Division mit dem ersten Faktor der rechten Seite in (55) die Beziehung

$$\frac{4\pi^2 + \omega^2(\varrho)}{\pi \sum_{\mu=1}^{m} (\varrho \cdot \lambda_\mu(\varrho) + \varkappa_\mu(\varrho))} \leq \sum_{\mu=1}^{m} \int_{B_\varrho^\mu} \left| \frac{dv}{d\zeta_\mu} \right|^2 \cdot \frac{1}{D_\mu(\zeta_\mu)} |d\zeta_\mu|.$$

Diese Ungleichung integrieren wir nach ϱ zwischen den Grenzen ϱ', ϱ'' $(\varrho' < \varrho'')$ und beachten, daß die rechte Seite dann höchstens gleich dem Inhalt der von $B_{\varrho,v}$ überstrichenen Fläche ist, der nicht größer ist als $2\pi(\log u_2(\varrho'') - \log u_1(\varrho'))$:

$$\int_{\varrho'}^{\varrho''} \frac{4\pi^2 + \omega^2}{\pi \sum \varrho \cdot \lambda_\mu + \varkappa_\mu} d\varrho \leq 2\pi \log \frac{u_2(\varrho'')}{u_1(\varrho')}.$$

Eine Umformung, die die $\omega(\varrho)$ enthaltenden Glieder auf eine Seite bringt, ergibt

$$\int_{\varrho'}^{\varrho''} \frac{2}{\varrho \cdot \sum\left(\lambda_\mu + \frac{\varkappa_\mu}{\varrho}\right)} d\varrho - \log \frac{u_1(\varrho'')}{u_1(\varrho')} \leq \omega(\varrho'') - \frac{1}{2\pi^2} \int_{\varrho'}^{\varrho''} \frac{\omega^2}{\varrho \cdot \sum\left(\lambda_\mu + \frac{\varkappa_\mu}{\varrho}\right)} d\varrho .$$

Wir definieren eine Menge P in folgender Weise. Es sei

$$\varrho \in \mathsf{P}, \text{ wenn } \omega(\varrho) - \frac{1}{2\pi^2} \int_{\varrho'}^{\varrho} \frac{\omega^2}{\varrho \cdot \sum\left(\lambda_\mu + \frac{\varkappa_\mu}{\varrho}\right)} d\varrho > 1 \text{ und } \varrho \geq \varrho' .$$

Setzen wir

$$\alpha(\varrho) = \frac{1}{2\pi^2} \int_{\varrho'}^{\varrho} \frac{\omega^2}{\varrho \cdot \sum\left(\lambda_\mu + \frac{\varkappa_\mu}{\varrho}\right)} d\varrho ,$$

so ist für $\varrho \in \mathsf{P}$

$$\frac{d\alpha}{d\varrho} = \frac{1}{2\pi^2} \cdot \frac{\omega^2}{\varrho \cdot \sum\left(\lambda_\mu + \frac{\varkappa_\mu}{\varrho}\right)} \geq \frac{1}{2\pi^2} \cdot \frac{(1+\alpha)^2}{\varrho \cdot \sum\left(\lambda_\mu + \frac{\varkappa_\mu}{\varrho}\right)} .$$

Weil $\alpha(\varrho)$ eine monotone Funktion ist, ergibt die Integration über die Menge P die Ungleichung

$$\frac{1}{2\pi^2} \int_{\mathsf{P}} \frac{d\varrho}{\varrho \cdot \sum\left(\lambda_\mu + \frac{\varkappa_\mu}{\varrho}\right)} \leq \int_0^\infty \frac{d\alpha}{(1+\alpha)^2} < \infty .$$

Das Maß von P gemessen in der Metrik

$$df(\varrho) = \frac{d\varrho}{\varrho \cdot \sum_{\mu=1}^{m} \left(\lambda_\mu(\varrho) + \frac{\varkappa_\mu(\varrho)}{\varrho}\right)} . \tag{56}$$

ist daher endlich. Weil $\sum\left(\lambda_\mu + \frac{\varkappa_\mu}{\varrho}\right)$ beschränkt (und positiv) ist, gibt es also beliebig große Werte von ϱ'', für die die Ungleichung

$$\int_{\varrho'}^{\varrho''} \frac{2}{\varrho \cdot \sum\left(\lambda_\mu + \frac{\varkappa_\mu}{\varrho}\right)} d\varrho - \log \frac{u_1(\varrho'')}{u_1(\varrho')} \leq 1 \tag{57}$$

besteht. [$u_1(\varrho)$ ist daher nicht beschränkt, nach der Definition von $u_1(\varrho)$ folgt wieder Satz 1.] Die mit (57) gleichbedeutende Formel

$$u_1(\varrho) \geq C\, e^{f(\varrho)} \text{ mit } C > 0, \quad f(\varrho) = \int_{\varrho'}^{\varrho} df(\varrho) = \int_{\varrho'}^{\varrho} \frac{d\varrho}{\varrho \cdot \sum_{\mu=1}^{m} \left(\lambda_\mu(\varrho) + \frac{\varkappa_\mu(\varrho)}{\varrho}\right)} \tag{57'}$$

gilt bis auf eine Menge P von endlichem $f(\varrho)$-Maß. Aus Stetigkeitsgründen ist P eine Summe von offenen Intervallen.

Mit $\nu(\varrho) = O(\varrho)$ folgt aus (57') zunächst

$$1 \geq \overline{\lim} \frac{\log \nu(\varrho + 0)}{\log \varrho} \geq \mathrm{LIM} \frac{\log \nu(\varrho + 0)}{\log u_1(\varrho)} \cdot \frac{f(\varrho)}{\log \varrho};$$

das Zeichen LIM besage dabei: Limesbildung unter Weglassung der Ausnahmemenge P. Wir brauchen aber bei der Limesbildung die Ausnahmemenge P nicht auszuschließen, wir können das Zeichen LIM durch das Zeichen lim ersetzen: für $\varrho \in \mathsf{P}$ bezeichne δ_ϱ den Abstand des Punktes ϱ vom linken Endpunkt des Ausnahmeintervalles, in dem ϱ liegt; für $\varrho \notin \mathsf{P}$ sei $\delta_\varrho = 0$. Es ist dann stets

$$\frac{\log \varrho + O(1)}{f(\varrho - \delta_\varrho) + O(1)} \geq \frac{\log \nu(\varrho + 0)}{\log u_1(\varrho)}.$$

Es ist aber $f(\varrho) - f(\varrho - \delta_\varrho)$ beschränkt, also gilt auch

$$\overline{\lim} \frac{\log \nu(\varrho + 0)}{\log u_1(\varrho)} \leq \overline{\lim} \frac{\log \varrho}{f(\varrho)}. \tag{58}$$

Mit der Umkehrfunktion $\varrho_1(r)$ von $u_1(\varrho)$ ist nach (24), (30)

$$\frac{\log n(r)}{\log r} \leq \frac{\log \nu(\varrho_1(r) + 0)}{\log u_1(\varrho_1(r) - 0)},$$

also wegen der Stetigkeit von $\varrho_1(r)$ nach (58)

$$\overline{\lim} \frac{\log n(r)}{\log r} \leq \overline{\lim} \frac{\log \varrho}{f(\varrho)}.$$

Auf ähnliche Weise folgt mit positivem s

$$\overline{\lim} \frac{n(r)}{r^s} \leq C_0 \cdot \overline{\lim} \frac{\varrho}{e^{s f(\varrho)}}, \quad \text{wo} \quad C_0 > 0.$$

Nach den bekannten Ungleichungen

$$\overline{\lim} \frac{\log n(r)}{\log r} \geq \overline{\lim} \frac{\log N(r)}{\log r}; \quad \overline{\lim} \frac{n(r)}{r^s} \geq \overline{\lim} \frac{N(r)}{r^s}$$

folgt schließlich der

Satz 10. *$w(z)$ erzeuge eine* RIEMANNsche *Fläche $W \in \mathfrak{W}$. Es sei $f(\varrho) = \int_{\varrho_0}^{\varrho} d f(\varrho)$, wo $d f(\varrho)$ durch (56) bestimmt sei und ϱ_0 ($< \varrho$) eine hinreichend große feste Zahl sei. Dann gilt für die (obere) Wachstumsordnung s der Funktion $w(z)$ die Ungleichung*

$$s \leq \overline{\lim} \frac{\log \varrho}{f(\varrho)}.$$

Der Wachstumstypus der Oranung s ist höchstens maximal, mittel oder minimal, je nachdem

$$\overline{\lim} \frac{\varrho}{e^{s f(\varrho)}}$$

unendlich, positiv endlich oder Null ist.

Als Corollar folgt Satz 3, denn aus $\lim \sum_{\mu=1}^{m} \lambda_\mu(\varrho) = 0$ folgt $s \leq 0$, also $s = 0$.

Schlußbemerkung. Welche Ergebnisse wird man erwarten können, wenn man, an Stelle endlich vieler, unendlich viele Enden zuläßt? Hier müßten wohl gewisse Symmetriebedingungen erfüllt sein, die die Schwierigkeit unendlich vieler Enden mindern. Man könnte etwa von der $\wp$-Funktion ausgehen und durch Verschiebung der zweiblättrigen Windungspunkte unendlich viele mittelbare Randstellen der Konvergenzstärke p_ν hervorrufen. Als Anhaltspunkt für die Wachstumsordnung wird die Größe $2 + \frac{1}{\pi} \sum_{1}^{\infty} \operatorname{arc\,tg} \frac{p_\nu}{\pi}$ dienen können. Liegen die mittelbaren Randstellen so, daß ihre „Defekte" nicht vom starken Wachstum der Funktion aufgesogen werden[44], so haben die Randstellenkoordinaten Defekte; man wird annehmen, daß man auf diese Weise auch für unendlich viele verschiedene Stellen positive Defekte erzwingen kann; und es wäre denkbar, daß die Wachstumsordnung einer Funktion mit unendlich vielen Sorten positiven Defekts sich gar nicht von 2 allzusehr zu unterscheiden braucht, wenn die Defektsumme auch klein ist.

Literatur

[1] Ahlfors, L.: Über eine in der neueren Wertverteilungstheorie betrachtete Klasse transzendenter Funktionen. Acta Math. **58**, 375—406 (1932). — [2] Elfving, G.: Über eine Klasse von Riemannschen Flächen und ihre Uniformisierung. Acta Soc. Sci. Fenn., Nova Ser. **2**, Nr. 3, 1—60 (1934). — [3] Fischer, W.: Über die Riemannschen Flächen der Gaussschen Ψ-Funktion und der Mittag-Lefflerschen E_α-Funktionen. Mitt. Math. Sem. Gießen, H. **37**, 1—33 (1949). — [4] Hällström, G. af: Eine quasikonforme Abbildung mit Anwendung auf die Wertverteilungslehre. Acta Acad. Aboensis **18**, 8, 1—16 (1952). — [5] Huckemann, F.: Verschmelzung von Randstellen Riemannscher Flächen. Mitt. Math. Sem. Gießen H. **41**, 1—36 (1952). — [6] Huckemann, F.: Bestimmung der Wertverteilung der Gammafunktion aus ihrer Riemannschen Fläche. Math. Z. **59**, 375—382 (1954). — [7] Huckemann, F.: Zur Darstellung Riemannscher Flächen durch Streckenkomplexe. Math. Z. **64**, 215—239 (1956). — [8] Kneser, H., u. E. Ullrich: Funktionentheorie. Naturforschung und Medizin in Deutschland 1939—1946. Fiat Review, Bd. I (Reine Math.), Teil I, Kap. XII. — [9] Iversen, F.: Recherches sur les fonctions inverses des fonctions méromorphes. Thèse. Helsingfors **1914**. — [10] Le-Van, Th.: Über das Umkehrproblem der Wertverteilungslehre. Comment. Math. Helv. **23**, 26—49 (1949). — [11] MacLane, G. R.: Concerning the uniformization of certain Riemann Surfaces allied to the inverse-cosine and inverse-gamma surfaces. Trans. Amer. Math. Soc. **62**, 99—113 (1947). — [12] Nevanlinna, R.: Über die Herstellung transzendenter Funktionen als Grenzwerte rationaler Funktionen. Acta Math. **55**, 259—276 (1930). — [13] Nevanlinna, R.: Über Riemannsche Flächen mit endlich vielen Windungspunkten. Acta Math. **58**, 295—373 (1932). — [14] Nevanlinna, R.: Über die Riemannsche Fläche einer analytischen Funktion. Verh. Internat. Math. Kongr. Zürich 1932, **1**, S. 221—239. — [15] Nevanlinna, R.: Eindeutige analytische Funktionen, 2. Aufl. Berlin-Göttingen-Heidelberg 1953. — [16] Pöschl, K.: Über die Wertverteilung der erzeugenden Funktionen Riemannscher Flächen mit endlich vielen periodischen Enden. Math. Ann. **123**, 79—95 (1951). — [17] Pratje, I.: Iteration der Joukowski-Abbildung

[44] Man wird die Randstellen nur über „wenige" Koordinaten legen, entsprechend dem Verfahren, mit dem man unendlich viele positive Verzweigungsindizes erhalten kann. Vgl. af Hällström [4], Huckemann [5].

und ihre Streckenkomplexe. Mitt. Math. Sem. Gießen, H. **48**, 1—53 (1954). — [*18*] TEICHMÜLLER, O.: Eine Umkehrung des zweiten Hauptsatzes der Wertverteilungslehre. Dtsch. Math. **2**, 96—107 (1937). — [*19*] TEICHMÜLLER, O.: Untersuchungen über konforme und quasikonforme Abbildung. Dtsch. Math. **3**, 621—678 (1938). — [*20*] TEICHMÜLLER, O.: Vermutungen und Sätze über die Wertverteilung gebrochener Funktionen endlicher Ordnung. Dtsch. Math. **4**, 163—190 (1939). — [*21*] ULLRICH, E.: Über die Ableitung einer meromorphen Funktion. Sitzgsber. preuß. Akad. Wiss., phys.-math. Kl. **1929**. — [*22*] ULLRICH, E.: Über eine Anwendung des Verzerrungssatzes auf meromorphe Funktionen. J. reine angew. Math. **166**, 220—234 (1932). — [*23*] ULLRICH, E.: Zum Umkehrproblem der Wertverteilungslehre. Nachr. Ges. Wiss. Göttingen, math.-phys. Kl. I, N. F. **1**, Nr. 9, 135—150 (1936). — [*24*] ULLRICH, E.: Flächenbau und Wachstumsordnung bei gebrochenen Funktionen. Jber. dtsch. Math.-Verein. **46**, 232—274 (1936). — [*25*] ULLRICH, E.: Flächenbau und Wertverteilung. 9. Congr. Math. scand. Helsingfors, S. 179 bis 200, 1938. — [*26*] ULRICH, F. E.: The problem of type for a certain class of RIEMANN Surfaces. Duke Math. J. **5**, 567—589 (1939). — [*27*] WAGNER, H.: Über eine Klasse RIEMANNscher Flächen mit endlich vielen nur logarithmischen Windungspunkten. J. reine angew. Math. **175**, 6—49 (1936). — [*28*] WITTICH, H.: Über die Wachstumsordnung einer ganzen transzendenten Funktion. Math. Z. **51**, 1—16 (1949).

Gießen, Mathematisches Institut der Universität

(Eingegangen am 3. September 1955)

VOGEL, W.
Math. Zeitschr. Bd. 65, S. 283—308 (1956)

Die Annäherung guter Strategien bei einer gewissen Klasse von Spielen

Von
WALTER VOGEL

§ 1. Einleitung

Die vorliegende Arbeit befaßt sich mit einer naheliegenden und mancher Anwendungen fähigen Klasse von Zwei-Personen-Nullsummen-Spielen auf dem Einheitsquadrat. Ein Zwei-Personen-Nullsummen-Spiel auf dem Einheitsquadrat (s. z.B. MCKINSEY [9]) kann folgendermaßen erklärt werden: Gegeben ist eine Funktion $K(x; y)$, der Kern des Spieles, wobei $0 \leq x; y \leq 1$. Jeder der beiden Spieler wählt eine Verteilungsfunktion über der abgeschlossenen Strecke $[0; 1]$; diese Verteilungsfunktionen, sie sollen F und G heißen, werden Strategien genannt. Eine Strategie ist also eine Verteilungsfunktion, welche nur das Intervall $[0; 1]$ mit Gewicht belegt, d.h. eine Funktion F mit folgenden Eigenschaften: F sei von rechts stetig und monoton steigend, ferner gelte $F(x) = 0$ für $x < 0$ und $F(x) = 1$ für $x \geq 1$. Der Ausdruck $M(F; G) = \int\int K(x; y)\, dF(x)\, dG(y)$ gibt dann die Zahlung an, welche der zweite Spieler an den ersten zu leisten hat. Die Aufgabe des ersten Spielers besteht darin, eine Strategie F zu wählen, welche ihm ein möglichst günstiges Ergebnis verspricht. Dabei wird vorausgesetzt, daß er nicht weiß, welche Strategie sein Gegenspieler wählt.

Machen wir keine weiteren Voraussetzungen über $K(x; y)$, so gilt

$$\sup_F \inf_G M(F; G) \leq \inf_G \sup_F M(F; G). \tag{1}$$

Wir wollen annehmen, $K(x; y)$ sei so beschaffen, daß in (1) das Gleichheitszeichen steht:

$$\sup_F \inf_G M(F; G) = \inf_G \sup_F M(F; G) = v. \tag{2}$$

v heißt der Wert des Spieles. Offenbar gibt es dann zu vorgegebenem $\varepsilon > 0$ stets Strategien F_ε und G_ε derart, daß

$$\int K(x; y)\, dF_\varepsilon(x) \geq v - \varepsilon \quad \text{und} \quad \int K(x; y)\, dG_\varepsilon(y) \leq v + \varepsilon$$

gilt. Solche Strategien heißen ε-gute Strategien. Benutzt der erste Spieler eine ε-gute Strategie F_ε, so gilt für beliebiges G die Ungleichung $M(F_\varepsilon; G) \geq v - \varepsilon$. Er kann sich also durch Wahl von F_ε die Auszahlung $v - \varepsilon$ sichern, unabhängig davon, welche Strategie der zweite Spieler wählt.

Es kann nun sein, daß in dem Ausdruck (2) das Supremum und das Infimum angenommen werden, also $\max_F \min_G M(F; G) = \min_G \max_F M(F; G) = v$.

In diesem Falle gibt es Strategien F und G derart, daß

$$\int K(x;y)\,dF(x) \geq v \quad \text{und} \quad \int K(x;y)\,dG(y) \leq v$$

gilt. Solche Strategien heißen gute Strategien.

In § 2 geben wir ein Kriterium dafür an, daß eine Folge ε-guter Strategien gegen eine gute Strategie konvergiert. Dabei wird vorausgesetzt, daß der Kern des Spieles von beschränkter Schwankung sei.

Die folgenden Paragraphen dienen dazu, dieses Kriterium auf Spiele anzuwenden, deren Kern die Form

$$K(x;y) = f(x) - p\,g(y) \begin{cases} +p\,f(x)\,g(y) & \text{für } x < y \\ +0 & \text{für } x = y \\ -f(x)\,g(y) & \text{für } x > y \end{cases} \tag{3}$$

hat. Dabei seien f und g halbstetig nach oben, $0 \leq f;\ g \leq 1$ und $p \geq 0$.

Wir werden in § 6 zeigen, daß wir ohne Beschränkung der Allgemeinheit f und g monoton wachsend annehmen können. Dann sind f und g von rechts stetig und der Kern ist von beschränkter Schwankung.

Der Spieltyp mit monotonem f und g gehört in die Klasse der Zeitspiele, welche folgende Deutung zulassen: x bzw. y sind die Zeitpunkte, in welchen der erste bzw. zweite Spieler handeln. Für jeden Spieler ist es vorteilhaft, so spät wie möglich zu handeln, vorausgesetzt, daß es ihm gelingt, seinem Gegner zuvorzukommen. SHIFFMAN [12] und KARLIN [8] haben bewiesen, daß es bei Zeitspielen stets gute Strategien für beide Spieler gibt, wenn man voraussetzt, daß der Kern außerhalb der Hauptdiagonalen zweimal stetig differenzierbar und streng monoton ist. Wir kommen hier mit wesentlich schwächeren Voraussetzungen aus.

Das Spiel mit dem oben angegebenen Kern (3) heiße ein Duell. Wir zeigen nun zwei Anwendungen des Duells. Die erste dient gleichzeitig zur Erklärung des Namens.

(i) Jeder Spieler hat einen gewissen Munitionsvorrat, den er im Zeitintervall $[0;1]$ verschießen muß. Die Verteilung der Schußdichte über dieses Intervall bleibt seiner Willkür überlassen und bildet seine Strategie. Wir nennen die beiden Spieler A und B. Verschießt A seinen gesamten Munitionsvorrat zur Zeit x, so sei seine Treffwahrscheinlichkeit $f(x)$. Verschießt B seinen gesamten Munitionsvorrat zur Zeit y, so sei seine Treffwahrscheinlichkeit $g(y)$. Dabei soll jeder Spieler nur dann zum Schuß kommen, wenn er noch nicht getroffen worden ist. Wenn also A zur Zeit x schießt und B zur Zeit y schießt, und wenn $x < y$ ist, so wird die Treffwahrscheinlichkeit für B gleich $\big(1-f(x)\big)\,g(y)$, während A die Treffwahrscheinlichkeit $f(x)$ hat. Bewertet A seinen Erfolg mit $+q$ und den Erfolg von B mit $-p$, so ist sein Erwartungswert

$$K(x;y) = q\,f(x) - p\big(1-f(x)\big)\,g(y) = q\,f(x) - p\,g(y) + p\,f(x)\,g(y),$$

wenn A zur Zeit x und B zur Zeit y schießt und außerdem $x<y$ ist. Entsprechend ergibt sich für $x>y$:

$$K(x;y)=q\big(1-g(y)\big)f(x)-pg(y)=qf(x)-pg(y)-qf(x)g(y).$$

Für $x=y$ wird der Erwartungswert für A: $K(x;y)=qf(x)-pg(y)$. Zusammengefaßt ist also

$$K(x;y)=qf(x)-pg(y)\begin{cases}+pf(x)g(y) & \text{für } x<y\\ +0 & \text{für } x=y\\ -qf(x)g(y) & \text{für } x>y.\end{cases}$$

Da es bei der Bewertung offenbar nur auf das Verhältnis von q und p ankommt, so setzen wir $q=1$, also

$$K(x;y)=f(x)-pg(y)\begin{cases}+pf(x)g(y) & \text{für } x<y\\ +0 & \text{für } x=y\\ -f(x)g(y) & \text{für } x>y.\end{cases}$$

Auf diese Weise ergibt sich der Kern (3).

Der zweite Spieler hat im allgemeinen eine andere Bewertung als der erste Spieler, so daß das Spiel kein Nullsummen-Spiel wird. Setzen wir aber voraus, daß die Spieler nicht die Möglichkeit haben, miteinander zu verhandeln, oder daß die Verhandlungen gescheitert sind, so ist es für den ersten Spieler naheliegend, so zu handeln, als ob er sich in einem Nullsummen-Spiel befände. Die Deutung als Duell legt eine solche Haltung besonders nahe.

Im allgemeinen werden die beiden Spieler nicht alle ihre Munition auf einmal verschießen, sondern sie irgendwie im Intervall $[0;1]$ verteilen. Diese Häufigkeitsverteilungen sind die „Strategien" der Spieler (s. die Definition von Strategie auf S. 283).

Eine andere Möglichkeit, ein Duell als ein Zwei-Personen-Nullsummen-Spiel zu behandeln, führen DANSKIN und GILLMAN in [4] aus.

(ii) Die zweite Anwendung behandelt das Problem der Werbung von zwei Kaufleuten (sie sollen A und B heißen) um einen Kunden. Der Widerstand des Kunden gegen die Werbung sei eine Funktion der Zeit. Wirbt A zur Zeit x, so sei die Kaufwahrscheinlichkeit des Kunden $f(x)$. Wirbt B zur Zeit y, so sei die Wahrscheinlichkeit dafür, daß der Kunde bei B kauft, gleich $g(y)$. Außerdem soll gelten: Wenn der Kunde schon bei A gekauft hat, so kauft er nicht mehr bei B und umgekehrt. A bewertet es mit $+1$, wenn der Kunde bei ihm kauft, und mit $-p$, wenn der Kunde bei B kauft. Der Erwartungswert für A, wenn er zur Zeit x wirbt und B zur Zeit y wirbt, ist dann wieder der Kern (3). Hier ist es freilich naheliegend, das Spiel nicht mehr als ein Nullsummen-Spiel zu behandeln, sondern anzunehmen, daß A und B den Kunden untereinander aushandeln und einer von beiden gegen Zahlung einer Kompensation an den anderen von diesem den Kunden überlassen bekommt. Zu dieser Anwendung vergleiche man BURGER [2] und GILLMAN [6].

Wir werden die verschiedenen Deutungsmöglichkeiten nicht weiter diskutieren, sondern uns im folgenden stets an die Deutung als Duell halten und eine entsprechende Terminologie (z.B. Treffwahrscheinlichkeiten für die Funktionen f und g) gebrauchen.

§ 2. Eine Bedingung für die Konvergenz ε-guter Strategien

Wir interessieren uns für die Frage: Unter welchen Bedingungen konvergieren ε-gute Strategien gegen eine gute Strategie? Durch diese Frage werden wir darauf geführt, nach einem Kriterium für die Kompaktheit im Raum der Strategien zu suchen, wenn als Konvergenzbegriff die gleichmäßige Konvergenz benutzt wird. Da sich das gesuchte Kriterium (der Satz 2 dieses Paragraphen) genau so leicht für Funktionen beschränkter Schwankung aussprechen läßt wie für Strategien, so werden wir es in dieser etwas allgemeineren Form beweisen. Wir führen zuerst alle Definitionen und Sätze an, dann folgen die Beweise.

Wir betrachten Funktionen auf dem Intervall $[0;1]$. Die Abkürzung v. b. Sch. bedeutet „von beschränkter Schwankung". Als Norm führen wir den Höchstbetrag ein: $\|k\| = \sup_{x \in [0,1]} |k(x)|$; dann bedeutet $\|k_n - k\| \to 0$ die gleichmäßige Konvergenz der Funktionenfolge $\{k_n\}$ gegen k. Wir schreiben dafür auch $k_n \Rightarrow k$. Eine Teilung $\mathfrak{T}$ ist eine endliche Punktmenge $\{x_i\}$: $0 = x_0 < x_1 < \cdots < x_i < \cdots < x_r = 1$.

Satz 1. *Zu jedem $k(x)$ v.b.Sch. und zu jedem $\varepsilon > 0$ gibt es eine Teilung $\mathfrak{T} = \mathfrak{T}(\varepsilon, k)$ mit der Eigenschaft $|k(x') - k(x'')| < \varepsilon$ für $x_i < x'$, $x'' < x_{i+1}$ und $i = 0, 1, \ldots, r-1$.*

Wir bezeichnen eine solche Teilung $\mathfrak{T}(\varepsilon, k)$ als eine ε-Teilung.

Definition. *Eine Menge $\mathfrak{K}$ von Funktionen k heiße* gleichgradig T-stetig, *wenn es zu jedem $\varepsilon > 0$ eine Teilung $\mathfrak{T}(\varepsilon)$ gibt der Art, daß für alle $k \in \mathfrak{K}$ die Ungleichung $|k(x') - k(x'')| < \varepsilon$ gilt, sobald x' und x'' zwischen denselben Nachbarpunkten der Teilung liegen.*

Satz 2. *Sei $\mathfrak{K}$ eine unendliche Menge von Funktionen v.b.Sch. auf dem Intervall $[0, 1]$. Es seien alle Funktionen aus $\mathfrak{K}$ und die Gesamtschwankung aller Funktionen aus $\mathfrak{K}$ gleichmäßig beschränkt. Wenn $\mathfrak{K}$ gleichgradig T-stetig ist, so gibt es in $\mathfrak{K}$ stets eine Folge $\{k_n\}$, welche gleichmäßig gegen eine Funktion k_0 v.b.Sch. konvergiert.*

Dieser Satz, welcher dem Theorem von ARZELA-ASCOLI über gleichgradig stetige Funktionen (s. NATANSON [*10*] S. 433) entspricht, stellt das gesuchte Kompaktheitskriterium dar.

Satz 3. *Wenn $k_n \Rightarrow k$ und sowohl k_n als auch k v.b.Sch. ist, so gilt: $\{k_n\}$ ist gleichgradig T-stetig.*

Anwendung auf Strategien. *Sei $\{F_n\}$ eine gleichgradig T-stetige Folge von Strategien. Dann gibt es eine Teilfolge $\{F_{n'}\}$ und eine Strategie F so, daß $F_{n'} \Rightarrow F$. Umgekehrt folgt aus $F_n \Rightarrow F$, daß $\{F_n\}$ gleichgradig T-stetig ist.*

Satz 4. *Für eine Folge* $\{F_n\}$ *von Strategien ist* $F_n \Rightarrow F$ *gleichwertig mit*

$$(1) \qquad \begin{cases} \lim F_n(x) & = F(x) \\ \lim F_n(x-0) & = F(x-0) \\ \lim F_n(x+0) & = F(x+0). \end{cases}$$

Einen Beweis hierzu findet man bei FRÉCHET [5], S. 274–276.

Satz 5. *Es gelte* $\lim \int k(x)\, dF_n(x) = \int k(x)\, dF(x)$ *für jede Funktion* k *v.b.Sch. und* F *und die* F_n *seien Strategien. Dann folgt:* $F_n \Rightarrow F$. *Umgekehrt folgt aus* $F_n \Rightarrow F$, *daß für jede Funktion* k *v.b.Sch. die Gleichung* $\lim \int k(x)\, dF_n(x) = \int k(x)\, dF(x)$ *gilt.*

Alle Integrale sind LEBESGUE-STIELTJES-Integrale. Vergleiche dazu etwa CRAMÉR [3] S. 33 u.f.

Anmerkung. Gewöhnlich betrachtet man den Fall, daß die Gleichung $\lim \int f(x)\, dF_n(x) = \int f(x)\, dF(x)$ für jede stetige Funktion $f(x)$ gilt. Dies ist genau dann der Fall, wenn die Gleichung $\lim F_n(x) = F(x)$ für jeden Stetigkeitspunkt der Verteilungsfunktion $F(x)$ erfüllt ist. Siehe dazu: CRAMÉR [3] S. 74, GNEDENKO-KOLMOGOROV [7] § 9, MCKINSEY [9] S. 177 und NATANSON [10] S. 237.

Wir werden nun die Sätze 1 bis 5 beweisen und sie anschließend auf die Frage der Konvergenz ε-guter Strategien gegen eine gute Strategie anwenden.

Beweis von Satz 1. Wir beweisen den Satz zunächst für monoton steigende Funktionen. Sei k monoton steigend, $k(0) = a$ und $k(1) = b$. Wir wählen eine ganze Zahl $n > 0$ so, daß $1/n < \varepsilon$ ist. Man bilde nun:

$$x_i = \inf\left\{x \mid k(x) \geq \frac{b-a}{n}\, i + a\right\}$$

für $i = 0, 1, 2, \ldots, n$. Es kann vorkommen, daß mehrere aufeinanderfolgende x_i einander gleich sind. Wir wollen sie dann als einen Punkt zählen. Wir werden zeigen, daß diese x_i eine Teilung $\mathfrak{T}(\varepsilon, k)$ mit den in Satz 1 behaupteten Eigenschaften bilden. Aus $x_i < x' < x_{i+1}$ folgt nämlich

$$a + \frac{b-a}{n}\, i \leq k(x') < a + \frac{b-a}{n}\, (i+1).$$

Genau dieselbe Ungleichung gilt für $k(x'')$, wenn $x_i < x'' < x_{i+1}$ ist. Dann ist aber $|k(x') - k(x'')| \leq 1/n < \varepsilon$. Also bilden die x_i eine ε-Teilung.

Ist k monoton fallend, so ist offenbar die ε-Teilung der monoton steigenden Funktion $-k$ eine ε-Teilung für k.

Sei nun k v.b.Sch., also als Differenz zweier monotoner Funktionen p und q darstellbar: $k(x) = p(x) - q(x)$.

Wir bilden $\mathfrak{T}(\varepsilon, k) = \mathfrak{T}(\varepsilon/2, p) \cup \mathfrak{T}(\varepsilon/2, q)$, also diejenige Teilung, welche sowohl die Teilpunkte einer $\varepsilon/2$-Teilung von p als auch die Teilpunkte einer $\varepsilon/2$-Teilung von q enthält. Sind nun x_i und x_{i+1} aufeinanderfolgende Teilpunkte der Teilung $\mathfrak{T}(\varepsilon, k)$, dann gilt für $x_i < x'$, $x'' < x_{i+1}$ sowohl

$$|p(x') - p(x'')| < \varepsilon/2 \quad \text{als auch} \quad |q(x') - q(x'')| < \varepsilon/2.$$

Daraus folgt aber:

$$|k(x') - k(x'')| = |p(x') - p(x'') - q(x') + q(x'')| \leq |p(x') - p(x'')| + \\ + |q(x') - q(x'')| < \varepsilon.$$

Also ist $\mathfrak{T}(\varepsilon, k)$ eine ε-Teilung für die Funktion $k(x)$, und der Satz ist bewiesen.

Beweis von Satz 2. Nach dem Auswahlsatz von HELLY (s. NATANSON [*10*] S. 225) können wir aus der Menge $\mathfrak{R}$ eine Folge $\{k_n\}$ auswählen, welche in jedem Punkte des Intervalls $[0; 1]$ gegen eine Funktion k_0 v. b. Sch. konvergiert. Zum Beweise unseres Satzes brauchen wir also nur noch zu zeigen, daß diese Folge gleichmäßig gegen k_0 konvergiert.

Auf Grund der Voraussetzungen gibt es zu jedem $\varepsilon > 0$ eine Teilung $0 = x_0 < x_1 < \cdots < x_i < x_{i+1} < \cdots x_r = 1$ so, daß für alle n die Ungleichung $|k_n(x') - k_n(x'')| < \varepsilon/3$ gilt, wenn nur $x_i < x', x'' < x_{i+1}$ ist. Setzen wir zur Abkürzung $\frac{1}{2}(x_i + x_{i+1}) = y_i$, so folgen die Gleichungen $\lim k_n(x_i) = k_0(x_i)$ (für $i = 0, 1, 2, \ldots, r$) und $\lim k_n(y_i) = k_0(y_i)$ ($i = 0, 1, \ldots, r-1$). Daher gibt es ein $N = N(\varepsilon)$ mit der Eigenschaft, daß für alle $n, m > N(\varepsilon)$ folgende Ungleichungen gelten: $|k_n(x_i) - k_m(x_i)| < \varepsilon$ (für $i = 0, 1, \ldots, r$) und $|k_n(y_i) - k_m(y_i)| < \varepsilon/3$ (für $i = 0, 1, \ldots, r-1$). Sei nun x beliebig im Intervall $[0; 1]$ vorgegeben. Wir wollen zeigen, daß $|k_n(x) - k_m(x)| < \varepsilon$ ist. Ist $x = x_i$, so sind wir fertig! Ist $x_i < x < x_{i+1}$, so gilt (für $n, m > N(\varepsilon)$): $|k_n(x) - k_m(x)| \leq |k_n(x) - k_n(y_i)| + |k_n(y_i) - k_m(y_i)| + |k_m(y_i) - k_m(x)| < \varepsilon$. Läßt man nun m nach Unendlich gehen, so folgt für $n > N(\varepsilon)$ die Ungleichung $|k_n(x) - k_0(x)| \leq \varepsilon$, und zwar für jedes $x \in [0; 1]$. Damit ist der Satz 2 bewiesen.

Beweis von Satz 3. Da $k_n \Rightarrow k$ gilt, so gibt es zu jedem $\varepsilon > 0$ ein $N = N(\varepsilon)$ der Art, daß $|k_n(x) - k(x)| < \varepsilon/3$ ist, und zwar für alle x und für alle $n > N(\varepsilon)$. Da k nach Voraussetzung v. b. Sch. ist, so gibt es nach Satz 1 eine $\varepsilon/3$-Teilung $\mathfrak{T}(\varepsilon/3, k)$. Sind x_i und x_{i+1} aufeinanderfolgende Teilpunkte von $\mathfrak{T}(\varepsilon/3, k)$, so gilt für $x_i < x'; x'' < x_{i+1}$ die Ungleichung $|k(x') - k(x'')| < \varepsilon/3$. Wir zeigen, daß $\mathfrak{T}(\varepsilon/3, k)$ eine ε-Teilung für alle k_n mit $n > N(\varepsilon)$ ist. Es gilt nämlich für $x_i < x', x'' < x_{i+1}$ und $n > N(\varepsilon)$ die Abschätzung:

$$|k_n(x') - k_n(x'')| \leq |k_n(x') - k(x')| + |k(x') - k(x'')| + |k(x'') - k_n(x'')| < \varepsilon.$$

Folglich ist $\mathfrak{T}(\varepsilon/3, k)$ eine ε-Teilung für alle k_n mit $n > N(\varepsilon)$. Nehmen wir nun die Teilpunkte der endlich vielen Teilungen $\mathfrak{T}(\varepsilon, k_n)$ mit $n \leq N(\varepsilon)$ zu der Teilung $\mathfrak{T}(\varepsilon/3, k)$ hinzu, so erhalten wir eine ε-Teilung für alle k_n. Also ist $\{k_n\}$ gleichgradig T-stetig.

Zur Anwendung auf Strategien. Für jede Strategie F_n gilt: $F_n(x) = 0$ für $x < 0$ und $F_n(x) = 1$ für $x \geq 1$. Außerhalb des Intervalles $[0; 1]$ konvergiert also jede Folge von Strategien gleichmäßig. Da alle Strategien gleichmäßig beschränkt sind, und ihre Gesamtschwankungen gleich eins, also ebenfalls gleichmäßig beschränkt sind, so können wir für $x \in [0; 1]$ den Satz 2 anwenden. Es bleibt noch zu zeigen, daß die Grenzfunktion ebenfalls eine Strategie ist. Da alle F_n monoton steigen, so steigt auch F monoton. Ferner ist (auf Grund von Satz 4): $F(x+0) = \lim F_n(x+0) = \lim F_n(x) = F(x)$. $F(x)$ ist also ebenfalls eine Strategie.

Beweis von Satz 4. (Nach FRÉCHET [5] S. 274 bis 276.) 1. Es gelte (1). Wir wollen zeigen: $F_n \Rightarrow F$. Die Teilpunkte der $\varepsilon/2$-Teilung $\mathfrak{T}(\varepsilon/2, F)$ seien x_i $(i = 0, 1, 2, \ldots, r)$. Es gelten dann für $x_i < x < x_{i+1}$ die Ungleichungen:

$$F(x_{i+1} - 0) - \varepsilon/2 \leq F(x) \leq F(x_i + 0) + \varepsilon/2 = F(x_i) + \varepsilon/2.$$

Sei nun $N = N(\varepsilon)$ so gewählt, daß $|F_n(x_i) - F(x_i)| < \varepsilon/2$ ist (für $i = 0, 1, \ldots, r$), und daß $|F_n(x_i - 0) - F(x_i - 0)| < \varepsilon/2$ ist (für $i = 1, 2, \ldots, r$). Dann gelten für $x_i < x < x_{i+1}$ folgende Ungleichungen:

$$F_n(x) - F(x) \leq F_n(x_{i+1} - 0) - F(x) \leq F_n(x_{i+1} - 0) - F(x_{i+1} - 0) + \varepsilon/2 < \varepsilon$$

und

$$F_n(x) - F(x) \geq F_n(x_i) - F(x) \geq F_n(x_i) - F(x_i) - \varepsilon/2 > -\varepsilon.$$

Daraus folgt für $x_i < x < x_{i+1}$ die Ungleichung $|F_n(x) - F(x)| < \varepsilon$. Diese Ungleichung gilt aber auch für $x = x_i$. (Sie gilt natürlich auch, wenn x nicht im Intervall $[0; 1]$ liegt.) Damit ist bewiesen: Aus (1) folgt $F_n \Rightarrow F$.

2. Es gelte $F_n \Rightarrow F$. Dann gibt es zu jedem $\varepsilon > 0$ ein N so, daß die Ungleichung $|F_n(x) - F(x)| < \varepsilon$ für alle $n > N$ und für alle x erfüllt ist. Damit gilt auch für alle δ die Ungleichung $|F_n(x \pm \delta) - F(x \pm \delta)| < \varepsilon$. Dann gelten aber die drei Gln. (1), und der Satz 4 ist bewiesen.

Beweis von Satz 5. ERSTER TEIL: Aus

$$\lim \int k(x)\, dF_n(x) = \int k(x)\, dF(x) \tag{2}$$

für jede Funktion k v.b.Sch. folgt $F_n \Rightarrow F$. Sei

$$k_1(x) = \begin{cases} 1 & \text{für } x \leq \xi \\ 0 & \text{für } x > \xi \end{cases} \quad \text{und} \quad k_2(x) = \begin{cases} 1 & \text{für } x < \xi \\ 0 & \text{für } x \geq \xi. \end{cases}$$

Aus Gl. (2) folgt mit $k = k_1$ die Gleichung $\lim F_n(\xi) = F(\xi)$. Für $k = k_2$ ergibt sich $\lim F_n(\xi - 0) = F(\xi - 0)$. Ferner gilt: $F(x + 0) = F(x) = \lim F_n(x) = \lim F_n(x + 0)$. Die Gln. (1) aus Satz 4 sind erfüllt, also gilt $F_n \Rightarrow F$.

ZWEITER TEIL. Aus $F_n \Rightarrow F$, d.h. aus den Gln. (1) folgt (2). Es ist praktischer im nachfolgenden statt mit den Verteilungsfunktionen mit den zugehörigen Maßen zu arbeiten. μ_n sei das zur Strategie F_n gehörige Maß und μ gehöre zur Strategie F. Die erste und dritte Gleichung (1) lauten dann: $\lim \mu_n[0; x] = \mu[0; x]$ und $\lim \mu_n[0; x) = \mu[0; x)$. Zusammengefaßt folgt:

$$\lim \mu_n \mathfrak{J} = \mu \mathfrak{J}. \tag{3}$$

Dabei kann $\mathfrak{J}$ je nach Belieben ein Intervall der Form $[a; b]$, $[a; b)$, $(a; b]$ oder $(a; b)$ sein. Die zweite Gl. (1) sagt nur aus, daß $F(x)$ eine Strategie ist.

Sei nun $k(x)$ eine beliebige Funktion v.b.Sch. Wir wählen eine Teilung $\mathfrak{T}(\varepsilon/2, k)$: $0 = x_0 < x_1 < \cdots < x_r = 1$. Wir setzen: $S_0 = [x_0]$; $S_1 = (x_0, x_1)$; $S_2 = [x_2]$; $\ldots$ $S_{2\nu} = [x_\nu]$; $S_{2\nu+1} = (x_\nu, x_{\nu+1})$; $\ldots$ $S_{2r} = [x_r]$. Es gilt also: $|k(x') - k(x'')| < \varepsilon/2$ für $x', x'' \in S_i$. Sei nun $\bar{k}_i = \sup_{x \in S_i} k(x)$ und $\underline{k}_i = \inf_{x \in S_i} k(x)$. Dann folgt: $|\bar{k}_i - \underline{k}_i| \leq \varepsilon/2$.

Die S_i sind meßbare Mengen und bilden eine Zerlegung des Intervalles $[0;1]$, d.h. es ist $[0;1]=\bigcup_{i=0}^{2r} S_i$ und $S_i \cap S_k = \varnothing$ für $i \neq k$. Dann gilt:

$$\sum_{i=0}^{2r} \underline{k}_i \mu S_i \leq \int k(x)\, d\mu \leq \sum_{i=0}^{2r} \overline{k}_i \mu S_i$$

(s. CRAMÉR [3] S. 33–35).

Sei $K \geq |k(x)|$ eine Schranke für $k(x)$. Auf Grund der Gl. (3) kann man erreichen, daß

$$|\mu S_i - \mu_n S_i| < \frac{\varepsilon}{4r K}$$

gilt, und zwar für alle $n > N(i)$. Da wir nur endlichviele $N(i)$ haben, so können wir N einheitlich für alle i wählen. Es folgt dann:

$$\sum_{i=0}^{2r} |\mu S_i - \mu_n S_i| < \frac{\varepsilon}{2K}$$

für alle $n > N$.

Wir können jetzt die Ungleichungen

$$\sum \underline{k}_i \mu S_i = \underline{\sum} \leq \int k(x)\, d\mu \leq \overline{\sum} = \sum \overline{k}_i \mu S_i$$

und

$$\sum \underline{k}_i \mu_n S_i = \underline{\sum}_n \leq \int k(x)\, d\mu_n \leq \overline{\sum}_n = \sum \overline{k}_i \mu_n S_i$$

auswerten. Aus diesen Ungleichungen folgt nämlich:

$$\int k(x)\, d\mu - \int k(x)\, d\mu_n = \quad \Delta \leq \overline{\sum} - \underline{\sum}_n = A$$

und

$$\int k(x)\, d\mu_n - \int k(x)\, d\mu = -\Delta \leq \overline{\sum}_n - \sum = B,$$

also $|\Delta| \leq \max(A; B)$. Die Gl. (2), oder $|\Delta| \to 0$, ist bewiesen, wenn wir zeigen können, daß $A \to 0$ und $B \to 0$ gilt. Wir führen nur den Fall $A \to 0$ vor ($B \to 0$ folgt auf die gleiche Weise). Sei $n > N$. Dann ergibt sich:

$$\begin{aligned} A \leq |A| &= |\sum \overline{k}_i \mu S_i - \sum \underline{k}_i \mu S_i + \sum \underline{k}_i \mu S_i - \sum \underline{k}_i \mu_n S_i| \leq \\ &\leq \sum |\overline{k}_i - \underline{k}_i| \mu S_i + |\sum \underline{k}_i (\mu S_i - \mu_n S_i)| \leq \\ &\leq \varepsilon/2 + K \cdot \sum |\mu S_i - \mu_n S_i| < \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

Damit ist der Satz 5 bewiesen.

Wir geben zur Erläuterung ein Beispiel, in welchem eine der Bedingungen verletzt ist. Sei

$$F_n(x) = I_{\frac{1}{2}-1/n}(x) = \begin{cases} 0 & \text{für} \quad x < \frac{1}{2} - 1/n \\ 1 & \text{für} \quad x \geq \frac{1}{2} - 1/n. \end{cases}$$

Dann ist $\lim F_n(x) = F(x) = I_{\frac{1}{2}}(x)$, aber die Bedingung $\lim F_n(x-0) = F(x-0)$ ist für $x = \frac{1}{2}$ nicht erfüllt, denn $\lim\limits_{n\to\infty} F_n(\frac{1}{2}-0) = \lim\limits_{n\to\infty} \lim\limits_{\delta\to 0} F_n(\frac{1}{2}-\delta) = \lim\limits_{n\to\infty} 1 = 1$

und $F(\frac{1}{2}-0)=\lim\limits_{\delta\to 0}\lim\limits_{n\to\infty}F_n(\frac{1}{2}-\delta)=\lim\limits_{\delta\to 0}0=0$. Wählt man nun

$$k(x)=\begin{cases}1 & \text{für } x<\frac{1}{2}\\ 0 & \text{für } x\geq\frac{1}{2},\end{cases}\qquad \text{so gilt für alle } n$$

$$\int k(x)\,dF_n(x)=1 \quad \text{während} \quad \int k(x)\,dF(x)=0$$

ist.

Als Anwendung der Sätze dieses Paragraphen auf die Theorie der Zwei-Personen-Nullsummen-Spiele auf dem Einheitsquadrat beweisen wir nun den

Satz 6. *Gegeben sei ein Spiel, welches einen Wert v besitzt. Der Kern $K(x;y)$ des Spieles sei für festes y eine Funktion v.b.Sch. in x. Es sei ferner $\{F_\varepsilon\}$ (mit $\varepsilon\to 0$) eine Folge ε-guter Strategien des ersten Spielers. (Es ist also:*

$$\int K(x;y)\,dF_\varepsilon(x)\geq v-\varepsilon.)$$

Dann gilt: Wenn $\{F_\varepsilon\}$ gleichgradig T-stetig ist, so gibt es eine Teilfolge $\{F_{\varepsilon'}\}$ mit $\varepsilon'\to 0$, welche gegen eine gute Strategie des ersten Spielers konvergiert.

Beweis. Auf Grund von Satz 3 wissen wir, daß es in der Folge $\{F_\varepsilon\}$ eine Teilfolge $\{F_{\varepsilon'}\}$ mit $\varepsilon'\to 0$ gibt, so, daß $F_{\varepsilon'}\Rightarrow F$ gilt. Der Satz 5 liefert dann die Gleichung $\lim\limits_{\varepsilon'\to 0}\int K(x;y)\,dF_{\varepsilon'}(x)=\int K(x;y)\,dF(x)$ und da $\int K(x;y)\,dF_{\varepsilon'}(x)\geq v-\varepsilon'$ ist, so folgt $\int K(x;y)\,dF(x)\geq v$.
Damit ist gezeigt, daß $\lim\limits_{\varepsilon'\to 0}F_{\varepsilon'}(x)=F(x)$ eine gute Strategie für den ersten Spieler ist.

In den folgenden Paragraphen wollen wir untersuchen, für welche Klassen von Duellen (ein Duell ist in § 1 auf S. 284 erklärt) die Voraussetzungen des Satzes 6 erfüllt sind.

§ 3. Einige Sätze über Duelle

Wir möchten die Ergebnisse von § 2 auf Duelle anwenden. Bevor wir uns im nächsten Paragraphen mit ε-guten Strategien beschäftigen können, müssen wir einige Sätze über die Struktur von Duellen herleiten.

Das Duell sei gegeben durch seinen Kern $K(x;y)$:

$$(1)\qquad K(x;y)=f(x)-p\,g(y)\begin{cases}+p\,f(x)\,g(y) & \text{für } x<y\\ +0 & \text{für } x=y\\ -f(x)\,g(y) & \text{für } x>y.\end{cases}$$

Dabei gelte: $0\leq f;g\leq 1$; $p\geq 0$; f und g seien monoton steigend und von rechts stetig. Wir sprechen dann von einem monotonen Duell. Diese Voraussetzungen sollen für den ganzen § 3 gelten.

Wir setzen nun fest:

$$\int_{[0,a)}^{a} f\,dF=\int_{[0,a)} f\,dF,\quad \int f\,dF=\int_{[0,1]} f\,dF.$$

Strategien für den ersten Spieler bezeichnen wir stets mit F, Strategien für den zweiten Spieler mit G. Es sei

$$U(F;y) = \int K(x;y)\,dF(x) = \int f\,dF - g(y)\{p + \int f\,dF + (1+p)\cdot\int_y f\,dF + f(y)\,\sigma_F(y)\};$$

$$V(x;G) = \int K(x;y)\,dG(y) = -p\int g\,dG + f(x)\{1 + p\int g\,dG - (1+p)\cdot\int_x g\,dG - p\,g(x)\,\sigma_G(x)\}.$$

Dabei sind $\sigma_F(x) = F(x) - F(x-0)$ und $\sigma_G(y) = G(y) - G(y-0)$ die Sprünge der Funktionen F und G.

Es gelten nun folgende Sätze:

Satz 7. *Sei* $f(x) = 0$ *für* $x < a$ *und*

$$F^*(x) = \begin{cases} 0 & \text{für } x < a \\ F(x) & \text{für } x \geq a, \end{cases}$$

dann gilt: $U(F^*;y) \geq U(F;y)$.

Satz 8. *Sei* $f(x) = c$ *für* $a \leq x < b$ *und*

$$F^*(x) = \begin{cases} F(x) & \text{für } x < a \\ F(b-0) & \text{für } a \leq x < b \\ F(x) & \text{für } b \leq x, \end{cases}$$

dann gilt: $U(F^*;y) \geq U(F;y)$.

Diese zwei Sätze legen es nahe, eine „im wesentlichen vollständige Klasse" von Strategien einzuführen (essentially complete class: s. BLACKWELL-GIRSHICK [1] Def. 5.4.10, S. 137). Eine Klasse $\mathfrak{C}$ von Strategien heißt im wesentlichen vollständig, wenn es zu jeder Strategie F, welche nicht in $\mathfrak{C}$ liegt, eine Strategie F^* in $\mathfrak{C}$ gibt, welche wenigstens ebensogut wie die Strategie F ist, d.h. es soll gelten: $U(F^*;y) \geq U(F;y)$.

Definition. *Für ein monotones Duell gelte* $F \in \mathfrak{C}_1$ *genau dann, wenn* $F(x) = 0$ *für* $x < \sup\{\xi \mid f(\xi) = 0\}$ *ist und wenn ferner* $\int_{(a,b)} dF = 0$ *gilt, falls* $f(x) = c$ *für* $a \leq x < b$ *ist.*

Es sollen also für die Strategien der Klasse $\mathfrak{C}_1$ die in den Sätzen 7 und 8 beschriebenen Reduktionen schon durchgeführt sein. $\mathfrak{C}_1$ ist eine im wesentlichen vollständige Klasse, so daß wir uns bei Fragen, welche gute oder ε-gute Strategien des ersten Spielers betreffen, stets auf die Betrachtung der Strategien der Klasse $\mathfrak{C}_1$ beschränken können.

Die Sätze 7 und 8 können wir auch beweisen, wenn darin die Rollen des ersten und zweiten Spielers vertauscht werden. Wir können daher für die Strategien des zweiten Spielers die im wesentlichen vollständige Klasse $\mathfrak{C}_2$ definieren.

Definition. *Für ein monotones Duell gelte* $G \in \mathfrak{C}_2$ *genau dann, wenn* $G(y) = 0$ *für* $y < \sup\{\eta \mid g(\eta) = 0\}$ *ist und wenn ferner* $\int_{(a,b)} dG = 0$ *gilt, falls* $g(y) = c$ *für* $a \leq y < b$ *ist.*

Beweis von Satz 7. $F^*(x)$ ist eine Strategie, welche das Intervall $[0; a)$ nicht mit Gewicht belegt. Für $y > a$ gilt:

$$\int_{}^{y} f\,dF^* = f(a)F(a) + \int_{(a,y)} f\,dF \text{ und } \int_{}^{y} f\,dF = f(a)\cdot\{F(a)-F(a-0)\} + \int_{(a,y)} f\,dF.$$

Also gilt für alle y die Ungleichung $\int_{}^{y} f\,dF^* \geq \int_{}^{y} f\,dF$. Genau so folgt

$$\int f\,dF^* \geq \int f\,dF \quad \text{und} \quad f(x)\,\sigma_{F^*}(x) \geq f(x)\,\sigma_F(x).$$

Nun ist aber

$$U(F; y) = (1-g(y))\int f\,dF - p\,g(y) + (1+p)\,g(y)\cdot\int_{}^{y} f\,dF + f(y)\,g(y)\,\sigma_F(y)$$

und

$$U(F^*; y) = (1-g(y))\int f\,dF^* - p\,g(y) + (1+p)\,g(y)\cdot\int_{}^{y} f\,dF^* + f(y)\,g(y)\,\sigma_{F^*}(y).$$

Da die Ausdrücke $\int f\,dF$, $\int_{}^{y} f\,dF$ und $f(y)\,\sigma_F(y)$ stets mit Faktoren auftreten, welche größer oder gleich Null sind, so folgt $U(F^*; y) \geq U(F; y)$. Damit ist der Satz bewiesen.

Beweis von Satz 8. $F^*(x)$ ist eine Strategie, welche das Intervall $(a; b)$ nicht mit Gewicht belegt. Es gilt:

$$\int f\,dF^* = \int_{}^{a} dF^* + f(a)\{F^*(b-0) - F^*(a-0)\} + \int_{[b,1]} f\,dF^* =$$

$$= \int_{}^{a} f\,dF + f(a)\{F(b-0) - F(a-0)\} + \int_{[b,1]} f\,dF = \int f\,dF.$$

Wir können jetzt den Ausdruck

$$U(F^*; y) = \int f\,dF - g(y)\{p + \int f\,dF\} + (1+p)\,g(y)\cdot\int_{}^{y} f\,dF^* + f(y)\,g(y)\,\sigma_{F^*}(y)$$

mit dem Ausdruck

$$U(F; y) = \int f\,dF - g(y)\{p + \int f\,dF\} + (1+p)\,g(y)\cdot\int_{}^{y} f\,dF + f(y)\,g(y)\,\sigma_F(y)$$

vergleichen. Es ist $U(F^*; y) \geq U(F; y)$ genau dann, wenn

$$(1+p)\cdot\int_{}^{y} f\,dF^* + f(y)\,\sigma_{F^*}(y) \geq (1+p)\cdot\int_{}^{y} f\,dF + f(y)\,\sigma_F(y)$$

ist.

Wir unterscheiden nun die drei Fälle $y \notin [a; b)$, $y \in (a; b)$ und $y = a$.

Für $y \notin [a; b)$ gilt offenbar in der letzten Ungleichung das Gleichheitszeichen.

Für $y \in (a; b)$ folgt:

$$(1+p)\cdot\int_{}^{y} f\,dF^* + f(y)\,\sigma_{F^*}(y) = (1+p)\cdot\int_{}^{a} f\,dF + (1+p)\,f(a)\int_{[a,b)} f\,dF$$

$$\geq (1+p)\cdot\int_{}^{y} f\,dF + f(y)\,\sigma_F(y).$$

Für $y=a$ folgt:

$$(1+p)\cdot\int^{a} f\,dF^{*}+f(a)\,\sigma_{F^{*}}(a)=(1+p)\cdot\int^{a} f\,dF+f(a)\int_{[a,b)}dF$$
$$\geq(1+p)\cdot\int^{a} f\,dF+f(a)\,\sigma_{F}(a).$$

Also ist $U(F^{*};y)\geq U(F;y)$ und der Satz ist bewiesen.

In Zukunft werden wir immer annehmen, daß unsere Strategien in den Klassen $\mathfrak{C}_1$ bzw. $\mathfrak{C}_2$ liegen.

Bei den folgenden Entwicklungen setzen wir voraus, daß es gute Strategien für beide Spieler geben soll. Wir möchten die Funktion g abändern, und zwar wie folgt:

$$g^{*}(y)=\begin{cases}0 & \text{für } y<a\\ g(y) & \text{für } y\geq a.\end{cases}$$

Dabei wollen wir versuchen, den Punkt a so zu wählen, daß gute Strategien in dem Duell mit g^{*} auch gute Strategien in dem Duell mit g sind. Außerdem soll für $y\in[a;1]$ gelten: $g^{*}(y)\geq\alpha>0$.

Satz 9. *Zu jedem monotonen Duell* mit $p\neq 0$ *kann man eine Zahl* $a\in[0;1]$ *und ein monotones Duell mit der Treffwahrscheinlichkeit*

$$g^{*}(y)=\begin{cases}0 & \textit{für } y<a\\ g(y) & \textit{für } y\geq a\end{cases}$$

für den zweiten Spieler angeben, so daß $g^{*}(y)\geq\alpha>0$ *für* $y\geq a$ *gilt und die guten Strategien (aus den Klassen* $\mathfrak{C}_1$ *und* $\mathfrak{C}_2$*) des Duells mit* g^{*} *auch gute Strategien des Duells mit* g *sind.*

Beweis. Sei $\sup\{y\,|\,g(y)=0\}=y_0$. Wenn $g(y_0)>0$ ist, so hat g schon die gewünschte Form. Wir nehmen daher an $g(y_0)=0$. Man bilde:

$$\sup\{g(y)\,(1-f(y))\}=s.$$

Zunächst zeigen wir, daß $s=0$ auf ein triviales Spiel führt. Aus $s=0$ folgt $f(x)=1$ für $x>y_0$. Da f von rechts stetig ist, so gilt auch $f(y_0)=1$. Ein monotones Duell mit $f(y_0)=1$ und $g(y_0)=0$ ist aber leicht zu lösen. $F(x)=I_{y_0}(x)$ ist eine gute Strategie für den ersten Spieler und es ist $U(I_{y_0};y)=1=v$ für alle y. Für den zweiten Spieler ist jede Strategie gut, da $K(x;y)\leq 1=v$ ist. Schließt man dieses Duell als uninteressant von den Betrachtungen aus, so darf man $s>0$ voraussetzen. Es gibt dann ein $y_1>y_0$, für das

$$g(y_1)\,(1-f(y_1))=c>0$$

ist. Da g von rechts stetig ist, so gibt es ein a, für das $y_0<a\leq y_1$ und $0=g(y_0)<g(a)\leq pc/(1+p)$ gilt. Mit einem so bestimmten a definieren wir nun:

$$g^{*}(y)=\begin{cases}0 & \text{für } y<a\\ g(y) & \text{für } y\geq a.\end{cases}$$

Seien nun F und G ein Paar guter Strategien für das Spiel mit g^*. Die Ausdrücke $U(F;y)$ und $V(x;G)$ sollen sich auf das gegebene Duell beziehen und die Ausdrücke $U^*(F;y)$ und $V^*(x;G)$ auf das neu konstruierte Duell. Es sei also: $U^*(F;y) \geq v^* \geq V^*(x;G)$. Wir wollen zeigen, daß $U(F;y) \geq v^* \geq V(x;G)$ gilt.

Es ist $V(x;G) = V^*(x;G)$, denn das Intervall, in welchem g und g^* voneinander verschieden sind, hat nach Voraussetzung (G sollte in $\mathfrak{C}_2$ liegen) das G-Maß Null. Ferner gilt $U(F;y) = U^*(F;y)$ für $y \geq a$, denn dann ist ja $g = g^*$. Wenn wir noch zeigen können, daß für $y < a$ die Ungleichung $U(F;y) \geq v^*$ gilt, so ist der Satz 9 bewiesen. Für $y < a$ folgt:

$$\begin{aligned} U(F;y) &= \int f\,dF - g(y)\{p + \int f\,dF - (1+p)\cdot\int^{y} f\,dF - f(y)\,\sigma_F(y)\} \\ &\geq \int f\,dF - g(a)\{1+p\} \geq \int f\,dF - p\,g(y_1)\{1 - f(y_1)\} \\ &\geq \int f\,dF - g(y_1)\{p - p\,f(y_1) + \int f\,dF - \int^{y_1} f\,dF - f(y_1)\,\sigma_F(y_1)\} \\ &\geq \int f\,dF - g(y_1)\{p + \int f\,dF - (1+p)\cdot\int^{y_1} f\,dF - f(y_1)\,\sigma_F(y_1)\} \\ &= U(F;y_1) = U^*(F;y_1) \geq v^*. \end{aligned}$$

Damit ist Satz 3 bewiesen.

Genau so zeigt man den

Satz 10. *Zu jedem monotonen Duell mit $p \neq 0$ und der Treffwahrscheinlichkeit f für den ersten Spieler kann man ein monotones Duell mit der Treffwahrscheinlichkeit f^* für den ersten Spieler angeben, so daß entweder $f^*(x) = 0$ oder $f^*(x) \geq \alpha > 0$ ist und die guten Strategien (aus den Klassen $\mathfrak{C}_1$ und $\mathfrak{C}_2$) des Duells mit f^* auch gute Strategien des Duells mit f sind.*

Anmerkung. Beim Beweise von Satz 10 (nicht aber beim Beweise von Satz 9) kann man die Voraussetzung $p \neq 0$ streichen.

Definition. *Ein α-Duell ist ein monotones Duell mit $f(x) = 0$ für $x < a_1$, $f(x) \geq \alpha_1 > 0$ für $x \geq a_1$, $g(y) = 0$ für $y < a_2$ und $g(y) \geq \alpha_2 > 0$ für $y \geq a_2$.*

Man kann die Sätze 9 und 10 zusammenfassen zu

Satz 11. *Zu jedem monotonen Duell mit $p \neq 0$ kann man ein α-Duell angeben der Art, daß die guten Strategien dieses α-Duells (aus den Klassen $\mathfrak{C}_1$ bzw. $\mathfrak{C}_2$) auch gute Strategien in dem gegebenen Duell sind.*

Beim Beweise von Satz 9 war

$$g(a) \leq \frac{p}{1+p}\, g(y_1)\,(1 - f(y_1)) \quad \text{mit} \quad y_1 \geq a. \tag{2}$$

Es gilt nun

Satz 12. *Sei $g(y) = 0$ für $y < a$ und*

$$F^*(x) = \begin{cases} 0 & \text{für} \quad x < a \\ F(x) & \text{für} \quad x \geq a. \end{cases}$$

Ferner gelte (2). Dann folgt aus $U(F;y) \geq v$ die Ungleichung $U(F^*;y) \geq v$, d.h. wenn F eine gute Strategie ist, so ist auch F^* eine gute Strategie.

Beweis. Es ist $\int f\,dF^* \geq \int f\,dF$ und, sofern nur $y > a$ ist, auch

$$\int_a^y f\,dF^* \geq \int_a^y f\,dF \quad \text{und} \quad \sigma_{F^*}(y) \geq \sigma_F(y).$$

Der Vergleich von $U(F^*; y)$ mit $U(F; y)$ liefert dann — genau so wie beim Beweise von Satz 7 auf S. 293 — für $y > a$ die Ungleichung $U(F^*; y) \geq U(F; y) \geq v$. Für $y < a$ folgt:

$$U(F^*; y) = \int f\,dF^* \geq \int f\,dF = U(F; y) \geq v.$$

Ist schließlich $y = a$, so folgt:

$$\begin{aligned} U(F^*; a) &= \int f\,dF^* - g(a)\{p + \int f\,dF^* - f(a)\,\sigma_{F^*}(a)\} \\ &\geq \int f\,dF^* - \frac{p}{1+p}\,g(y_1)\,(1 - f(y_1))\,\{1 + p\} \\ &\geq \int f\,dF - g(y_1)\{p - p f(y_1) + \int f\,dF - \int_a^{y_1} f\,dF - f(y_1)\,\sigma_F(y_1)\} \\ &\geq \int f\,dF - g(y_1)\{p + \int f\,dF - (1+p)\cdot \int_a^{y_1} f\,dF - f(y_1)\,\sigma_F(y_1)\} \\ &= U(F; y_1) \geq v. \end{aligned}$$

Damit ist Satz 12 bewiesen.

Mit Hilfe dieses Satzes ergibt sich der Beweis von

Satz 13. *Zu einem monotonen Duell mit $p \neq 0$ und* stetigem *g läßt sich ein α-Duell angeben der Art, daß eine gute Strategie F aus diesem α-Duell die Eigenschaft $\int_a dF = 0$ hat. Dabei ist $a = \sup\{x \mid f^*(x)\,g^*(x) = 0\}$. Außerdem ist F auch eine gute Strategie in dem gegebenen Duell.*

Beweis. Bei dem α-Duell sei $f^*(x) = 0$ für $x < a_1$, $f^*(x) \geq \alpha_1 > 0$ für $x \geq a_1$, $g^*(y) = 0$ für $y < a_2$ und $g^*(y) \geq \alpha_2 > 0$ für $y \geq a_2$. Ferner sei

$$a = \max(a_1; a_2) = \sup\{x \mid f^*(x)\,g^*(x) = 0\}.$$

Wir können ohne Beschränkung der Allgemeinheit $a > 0$ voraussetzen.

Wenn $a = a_1$ ist, so folgt $\int_a dF = 0$ aus Satz 7, S. 292. Wenn $a = a_2$ ist, so folgt aus dem Beweis von Satz 9 auf S. 294, daß die Ungleichung (2) gilt. Wir können also nach Satz 12 schließen, daß $\int_a dF = 0$ ist. Durch die Voraussetzung, daß g stetig sein soll, wird ausgeschlossen, daß man bei der Konstruktion des α-Duelles einfach $g^* = g$ setzen kann, denn dann brauchte ja die Ungleichung (2) nicht erfüllt zu sein.

Anmerkung. Die Sätze 9 bis 13 gelten auch dann noch, wenn man den Begriff „gut" durch den Begriff „ε-gut" ersetzt.

Die hier hergeleiteten Sätze werden wir zur Untersuchung der Frage, ob eine Menge von ε-guten Strategien gleichgradig T-stetig ist, gebrauchen. Vorher müssen wir uns aber überlegen, daß das Duell einen Wert hat und wie wir an ε-gute Strategien kommen können. Dazu dient der nächste Paragraph.

Ergänzung zu Satz 9. Beim Beweise von Satz 9 hatten wir ein Duell, für welches es ein x_0 mit der Eigenschaft $f(x_0)=1$ und $g(x_0)=0$ gab, von den weiteren Betrachtungen ausgeschlossen. Bei einem solchen Duell gilt natürlich der Satz 9 ebenfalls. Wie man auch $g(y)$ für $y>x_0$ abändert, in jedem Falle bleibt der Wert des Spieles $v=1$, und jede Strategie bleibt für den zweiten Spieler eine gute Strategie.

§ 4. Die Konstruktion ε-guter Strategien

Wir wollen zeigen; daß unser Duell einen Wert v hat, d.h. daß die Gleichung $\sup_F \inf_G M(F;G) = \inf_G \sup_F M(F;G)$ gilt. Der Beweis liefert gleichzeitig ein Konstruktionsverfahren für ε-gute Strategien.

Wir behandeln auch hier nur monotone Duelle, auf den allgemeinen Fall kommen wir in § 6 zurück.

Definition. *Ein* endliches Duell *ist ein endliches Zwei-Personen-Nullsummen-Spiel mit der Matrix* $(K_{\mu\nu})$:

$$K_{\mu\nu} = f_\mu - p\,g_\nu \begin{cases} +p\,f_\mu\,g_\nu & \text{für } \mu<\nu \\ +0 & \text{für } \mu=\nu \\ -f_\mu\,g_\nu & \text{für } \mu>\nu. \end{cases} \tag{1}$$

Dabei sei $\mu,\nu=0,1,\ldots,r$ *und* $0\leq f_\mu;\ g_\nu\leq 1$, *ferner* $p\geq 0$, $f_\mu\geq f_{\mu-1}$ *und* $g_\nu\geq g_{\nu-1}$.

Nach dem Hauptsatz der Theorie der endlichen Zwei-Personen-Nullsummen-Spiele (s. v. Neumann-Morgenstern [11] § 17.6, S. 153; McKinsey [9] 2, 3, S. 31; Blackwell-Girshick [1] 2.3.1, S. 41) hat das endliche Spiel einen Wert sowie gute Strategien für beide Spieler. Wir wollen ein endliches Duell betrachten, welches auf eine noch näher zu erläuternde Art das Duell auf dem Einheitsquadrat approximiert.

Gegeben sei ein monotones Duell durch seinen Kern $K(x;y)$. Der Kern soll den gleichen Voraussetzungen wie unter Formel (1) in § 3 auf S. 291 genügen. Es sollen also die Funktionen f und g von rechts stetig sein. Dann gibt es aber eine gemeinsame Teilung $\mathfrak{T}=\mathfrak{T}(\varepsilon)$ $(0=x_0<x_1<\cdots<x_\nu<\cdots<x_r=1)$ mit folgenden Eigenschaften: Aus $x_\nu\leq x<x_{\nu+1}$ folgt

$$|f(x_\nu)-f(x)|<\varepsilon \quad \text{und} \quad |g(x_\nu)-g(x)|<\varepsilon.$$

Wir setzen nun $f(x_\nu)=f_\nu$ und $g(x_\nu)=g_\nu$. Auf diese Weise erhalten wir ein endliches Duell mit der Matrix (1). Sein Wert sei $v^{(\varepsilon)}$, und es sei $\xi=(\xi_0\ldots\xi_r)$ eine gute Strategie für den ersten und $\eta=(\eta_0\ldots\eta_r)$ eine gute Strategie für den zweiten Spieler. Es gilt also:

$$\sum_{\mu=0}^{r} K_{\mu\nu}\,\xi_\mu \geq v^{(\varepsilon)} \quad \text{und} \quad \sum_{\nu=0}^{r} K_{\mu\nu}\,\eta_\nu \leq v^{(\varepsilon)}.$$

Wir definieren nun:

$$I_{x_\nu}(x) = \begin{cases} 0 & \text{für } x<x_\nu \\ 1 & \text{für } x\geq x_\nu, \end{cases}$$

und setzen $F_\varepsilon(x)=\sum_0^\nu \xi_\nu I_{x_\nu}(x)$. Für $y=x_\nu$ gilt dann:

$$U(F_\varepsilon;y)=\sum f_\mu\xi_\mu-g_\nu\{p+\sum f_\mu\xi_\mu\}+(1+p)\,g_\nu\sum_0^{\nu-1}f_\mu\xi_\mu+f_\nu g_\nu\xi_\nu=\sum K_{\mu\nu}\xi_\mu\geq v^{(\varepsilon)},$$

also:

$$U(F_\varepsilon;x_\nu)\geq v^{(\varepsilon)}. \tag{2}$$

Ferner wird für $y\in(x_\nu;x_{\nu+1})$:

$$U(F_\varepsilon;y)-\sum K_{\mu\nu}\xi_\mu=-\{g(y)-g_\nu\}\cdot\Big\{p+\sum f_\mu\xi_\mu-(1+p)\sum_0^\nu f_\mu\xi_\mu\Big\}+$$

$$+\,p f_\nu g_\nu\xi_\nu\geq-\{g(y)-g_\nu\}\cdot\Big\{p+\sum f_\mu\xi_\mu-(1+p)\sum_0^\nu f_\mu\xi_\mu\Big\}.$$

Berücksichtigt man nun:

$$0\leq\{g(y)-g_\nu\}\cdot\Big\{p+\sum f_\mu\xi_\mu-(1+p)\sum_0^\nu f_\mu\xi_\mu\Big\}\leq\{g(y)-g_\nu\}\cdot\{1+p\}\leq(1+p)\,\varepsilon,$$

so folgt:

$$U(F_\varepsilon;y)-\sum K_{\mu\nu}\xi_\mu\geq-(1+p)\,\varepsilon.$$

Zusammen mit $\sum K_{\mu\nu}\xi_\mu\geq v^{(\varepsilon)}$ ergibt sich dann:

$$U(F_\varepsilon;y)\geq v^{(\varepsilon)}-(1+p)\,\varepsilon.$$

Wegen (2) gilt damit für alle $y\in[0;1]$ die Ungleichung

$$U(F_\varepsilon;y)\geq v^{(\varepsilon)}-(1+p)\,\varepsilon. \tag{3}$$

Bildet man $G_\varepsilon(y)=\sum_0^\nu \eta_\nu I_{x_\nu}(y)$, so folgt durch eine entsprechende Rechnung:

$$V(x;G_\varepsilon)\leq v^{(\varepsilon)}+(1+p)\,\varepsilon. \tag{4}$$

Auf Grund der Ungleichungen (3) und (4) können wir nun den Beweis für die Gl. (2) aus § 1, also für

$$\sup_F\inf_G\iint K(x;y)\,dF(x)\,dG(y)=\inf_G\sup_F\iint K(x;y)\,dF(x)\,dG(y)$$

führen. Es ist

$$v_1=\sup_F\inf_G\iint K(x;y)\,dF(x)\,dG(y)\geq\inf_G\int U(F_\varepsilon;y)\,dG(y)$$
$$=\inf_y U(F_\varepsilon;y)\geq v^{(\varepsilon)}-(1+p)\,\varepsilon.$$

Daraus folgt: $v_1\geq v^{(\varepsilon)}-(1+p)\,\varepsilon$. Genau so ergibt sich:

$$v_2=\inf_G\sup_F\iint K(x;y)\,dF(x)\,dG(y)\leq v^{(\varepsilon)}+(1+p)\,\varepsilon.$$

Da immer $v_1\leq v_2$ gilt, so ist:

$$v^{(\varepsilon)}-(1+p)\,\varepsilon\leq v_1\leq v_2\leq v^{(\varepsilon)}+(1+p)\,\varepsilon. \tag{5}$$

Die Menge der $v^{(\varepsilon)}$ ist beschränkt, denn aus der Definition von $K(x;y)$ folgt, daß $-p\leq v^{(\varepsilon)}\leq+1$ gilt. Wir können daher eine Teilfolge herausgreifen der

Art, daß $\lim v^{(\varepsilon)} = v$ für $\varepsilon \to 0$ gilt. Aus (5) folgt dann: $v_1 = v_2 = v$. Dies ist aber die gewünschte Gleichung. Damit gilt:

Satz 14. *Das monotone Duell hat stets einen Wert.*

Wir haben bei diesem Beweise aber noch mehr erhalten. Aus (3) und (5) folgt $U(F_\varepsilon; y) \geq v - \varepsilon'$, wobei $\varepsilon' = 2(1+p)\varepsilon$ ist. Entsprechend läßt sich zeigen: $V(x; G_\varepsilon) \leq v + \varepsilon'$. Die guten Strategien eines geeignet gewählten endlichen Spieles haben uns also ε-gute Strategien für unser Duell geliefert. Diese ε-guten Strategien können wir stets so wählen, daß für sie gilt

Satz 15. *Die Strategie* $F_\varepsilon(x) = \sum_0^r \xi_\nu I_{x_\nu}(x)$ *hat folgende Eigenschaften: Aus* $f_\nu = 0$ *folgt* $\xi_\nu = 0$, *und aus* $f_\nu = f_{\nu-1}$ *folgt* $\xi_\nu = 0$.

Beweis. Auch für endliche Duelle gelten die Sätze 7 und 8 aus § 3 (S. 292), denn wir können ein endliches Duell auffassen als ein Duell auf dem Einheitsquadrat, bei welchem die Treffwahrscheinlichkeiten von rechts stetige Treppenfunktionen sind. Wir können daher die in den Sätzen 7 und 8 aus § 3 beschriebenen Reduktionen auch für die guten Strategien des endlichen Spiels durchführen. Dann besitzen aber die dazugehörigen Strategien F_ε gerade diejenigen Eigenschaften, welche in Satz 15 behauptet werden.

Wir können stets erreichen, daß $F_\varepsilon \in \mathfrak{C}_1$ liegt, doch müssen wir dazu die Teilung $\mathfrak{T}(\varepsilon)$, welche der Konstruktion von F_ε zugrunde liegt, noch etwas abändern. Wir müssen nämlich immer dann, wenn ein Teilpunkt in ein Konstanz-Intervall der Funktion f fällt, auch den Anfangspunkt dieses Konstanz-Intervalles mit zur Teilung $\mathfrak{T}$ rechnen. Diese Abänderung sei in Zukunft stets durchgeführt. Es gilt dann:

Satz 16. *Für die oben konstruierte ε-gute Strategie F_ε gilt: $F_\varepsilon \in \mathfrak{C}_1$.*

Anmerkung. Wir können einen dem Satz 15 entsprechenden Satz über $G_\varepsilon(y)$ angeben. Es ist aber nicht immer möglich, gleichzeitig die beiden Bedingungen $F_\varepsilon \in \mathfrak{C}_1$ und $G_\varepsilon \in \mathfrak{C}_2$ zu erfüllen. Es können Schwierigkeiten auftreten, wenn sowohl f als auch g unendlichviele Konstanz-Intervalle besitzen.

§ 5. Die Konvergenz ε-guter Strategien bei α-Duellen

Wir werden in diesem Paragraphen zeigen, daß bei gewissen Klassen von α-Duellen eine passend gewählte Folge von ε-guten Strategien gleichgradig T-stetig ist, also eine Teilfolge enthält, welche gegen eine gute Strategie konvergiert.

Gegeben sei ein monotones Duell mit $p \neq 0$. Wir wollen die in § 4 konstruierten ε-guten Strategien des ersten Spielers für $\varepsilon = 2(1+p)/n$ mit F_n bezeichnen. Es sei also $F_n(x) = \sum \xi_\nu I_{x_\nu}(x)$, wobei die Punkte x_ν eine Teilung $\mathfrak{T}_n$ bilden, so daß jedenfalls $g(x_\nu - 0) - g(x_{\nu-1}) < 1/n$ ist. Für F_n gelten die Sätze 15 und 16 aus § 4 (S. 299), d.h. F_n liegt in $\mathfrak{C}_1$. Wir wollen voraussetzen, daß $\mathfrak{T}_n \subseteq \mathfrak{T}_{n+1}$ gilt, d.h. die Punkte der Teilung $\mathfrak{T}_n$ sollen auch alle in der Teilung $\mathfrak{T}_{n+1}$ vorkommen.

Wir benutzen die folgenden Abkürzungen (dabei lassen wir den Index n zur Vereinfachung der Schreibweise überall weg). $\xi=(\xi_0,\ldots,\xi_r)$ und $\eta=(\eta_0,\ldots,\eta_r)$ seien gute Strategien für den ersten bzw. zweiten Spieler in dem endlichen Duell. Es sei

$$U_\nu(\xi)=\sum_{\mu=0}^{r}K_{\mu\nu}\xi_\mu \quad \text{und} \quad V_\mu(\eta)=\sum_{\nu=0}^{r}K_{\mu\nu}\eta_\nu.$$

Ferner setzen wir

$$C=\sum_{\mu=0}^{r}f_\mu\xi_\mu \quad \text{und} \quad D=\sum_{\nu=0}^{r}g_\nu\eta_\nu.$$

Dann ist

$$U_\nu(\xi)=C-g_\nu\Big\{p+C-(1+p)\sum_{0}^{\nu-1}f_\mu\xi_\mu-f_\nu\xi_\nu\Big\}$$

und

$$V_\mu(\eta)=-pD+f_\mu\Big\{1+pD-(1+p)\sum_{0}^{\mu-1}g_\nu\eta_\nu-pg_\mu\eta_\mu\Big\}.$$

Der Wert des endlichen Duells heiße v, also gilt: $U_\nu(\xi)\geq v\geq V_\mu(\eta)$. Im folgenden benutzen wir häufig den

Satz. *Wenn ξ und η gute Strategien sind, so folgt aus $\xi_\nu\neq 0$ die Gleichung $V_\nu(\eta)=v$ und aus $\eta_\mu\neq 0$ die Gleichung $U_\mu(\xi)=v$.*

Beweise hierfür findet man bei MCKINSEY [9] Theorem 2.12, S. 41; BLACKWELL-GIRSHICK [1] Theorem 2.7.1, S. 60 und bei v. NEUMANN-MORGENSTERN [11] 17:D, S. 161.

Definition. *Ein Duell heißt symmetrisch, wenn $f=g$ und $p=1$ ist.*

Wir behandeln zuerst das symmetrische α-Duell. Der Punkt

$$a=\sup\{x\,|\,f(x)=0\}$$

soll zu jeder Teilung $\mathfrak{T}_n$ gehören. Nach § 4 Satz 15 gilt dann $\xi_\nu=0$ für $x_\nu<a$. Sei nun $\xi_\nu\neq 0$ und $x_\nu>a$. Da das Duell symmetrisch ist, so gibt es auch eine gute Strategie für den zweiten Spieler mit $\eta_\nu\neq 0$. Dann ist aber $U_\nu(\xi)=v\leq U_{\nu-1}(\xi)$. Die Ungleichung $U_\nu(\xi)\leq U_{\nu-1}(\xi)$ lautet ausführlich:

$$C-f_\nu\Big\{1+C-2\sum_{0}^{\nu-1}f_\mu\xi_\mu\Big\}+f_\nu^2\xi_\nu\leq C-f_{\nu-1}\Big\{1+C-2\sum_{0}^{\nu-1}f_\mu\xi_\mu\Big\}-f_{\nu-1}^2\xi_{\nu-1}$$

oder

$$f_\nu^2\xi_\nu+f_{\nu-1}^2\xi_{\nu-1}\leq(f_\nu-f_{\nu-1})\Big\{1+C-2\sum_{0}^{\nu-1}f_\mu\xi_\mu\Big\}\leq 2(f_\nu-f_{\nu-1}).$$

Daraus folgt dann wegen $f_\nu^2\geq\alpha_1\alpha_2=\alpha>0$ die Ungleichung

$$\xi_\nu\leq\frac{2}{\alpha}(f_\nu-f_{\nu-1}) \quad \text{für} \quad x_\nu\neq a. \tag{1}$$

Wir können nun zeigen, daß $\{F_n\}$ gleichgradig T-stetig ist. Gegeben sei ein $\delta>0$. Man wähle N so, daß $\frac{2}{\alpha\cdot N}\leq\delta$ ist. Die Teilung $\mathfrak{T}_N$, d.h. die Teilung

$0=x_0<x_1<\cdots<x_\nu=1$ mit $f(x_k-0)-f(x_{k-1})<1/N$, ist eine δ-Teilung für alle F_n mit $n\geq N$. Dies folgt so: Für $n\geq N$ sind die Punkte von $\mathfrak{T}_N$ unter den Stützstellen von F_n enthalten. Es gilt daher für $x_k<x'<x''<x_{k+1}$ auf Grund von (1) die Abschätzung:

$$|F_n(x'')-F_n(x')|\leq\sum_{x'<x_\nu\leq x''}\xi_\nu\leq\sum_{x_k<x_\nu<x_{k+1}}\xi_\nu\leq\frac{2}{\alpha}\{f(x_{k+1}-0)-f(x_k)\}<\frac{2}{\alpha\cdot N}\leq\delta.$$

Nehmen wir nun zu der Teilung $\mathfrak{T}_N$ noch die endlichvielen Punkte hinzu, welche nötig sind, um eine δ-Teilung für die endlichvielen F_n mit $n<N$ zu erhalten, so haben wir eine δ-Teilung für $\{F_n\}$ konstruiert. Damit ist gezeigt: Für das symmetrische α-Duell ist $\{F_n\}$ gleichgradig T-stetig.

Als nächstes betrachten wir ein α-Duell, welchen aus einem monotonen Duell mit stetigem $g(y)$ und $p\neq 0$ hergeleitet worden ist. f und g seien die Treffwahrscheinlichkeiten im α-Duell. Wir setzen $a=\sup\{x\,|f(x)\,g(x)=0\}$. Auf Grund von Satz 13 in § 3 auf S. 296 dürfen wir dann für jede ε-gute Strategie F_ε voraussetzen, daß $\int_0^a dF_\varepsilon=0$ ist. Wir wollen die Strategien F_n so abändern, wie es in § 3 Satz 13 beschrieben wurde. Dann gilt $\xi_\nu=0$ für $x_\nu<a$, während die ξ_ν mit $x_\nu>a$ nicht geändert worden sind. Der Punkt a soll zu jeder Teilung $\mathfrak{T}_n$ gerechnet werden. Wir gehen genau so vor, wie bei der Behandlung des symmetrischen Duells. Zunächst versuchen wir ξ_ν für $a<x_\nu<1$ abzuschätzen. Sei $\xi_\nu\neq 0$ und $a<x_\nu<1$, dann folgt: $f_\nu g_\nu\geq\alpha_1\alpha_2=\alpha>0$. Wir müssen nun 3 Fälle unterscheiden:

1. Fall. $f_{\nu+1}>f_\nu$. Aus $\xi_\nu\neq 0$ folgt $V_\nu(\eta)=v\geq V_{\nu+1}(\eta)$. Die Ungleichung $V_\nu(\eta)\geq V_{\nu+1}(\eta)$ lautet ausführlich:

$$-pD+f_\nu\Big\{1+pD-(1+p)\sum_0^\nu g_\mu\eta_\mu\Big\}+f_\nu g_\nu\eta_\nu$$

$$\geq -pD+f_{\nu+1}\Big\{1+pD-(1+p)\sum_0^\nu g_\mu\eta_\mu\Big\}-pf_{\nu+1}g_{\nu+1}\eta_{\nu+1},$$

oder

$$f_\nu g_\nu\eta_\nu+pf_{\nu+1}g_{\nu+1}\eta_{\nu+1}\geq(f_{\nu+1}-f_\nu)\Big(1+pD-(1+p)\sum_0^\nu g_\mu\eta_\mu\Big).$$

Wir möchten hieraus schließen, daß $\eta_\nu\neq 0$ oder $\eta_{\nu+1}\neq 0$ ist. Wenn $\eta_\nu\neq 0$ ist, so brauchen wir nichts zu beweisen. Wenn $\eta_\nu=0$ ist, so müssen wir zeigen, daß $\eta_{\nu+1}\neq 0$ ist. Dies ist erreicht, wenn wir zeigen können, daß

$$A_\nu=1+pD-(1+p)\sum_0^\nu g_\mu\eta_\mu>0$$

ist, denn $f_{\nu+1}-f_\nu>0$ war vorausgesetzt worden. Den Beweis führen wir, indem wir aus folgenden drei Gleichungen einen Widerspruch herleiten:

$$A_{\nu_0}=0;\quad \eta_{\nu_0}=0\quad\text{und}\quad\xi_{\nu_0}\neq 0.$$

Da $g_\nu\leq 1$ ist, so folgt aus $A_{\nu_0}=0$ die Gleichung $1=D=\sum_0^{\nu_0}g_\mu\eta_\mu$. Es dürfen also nur den ν mit $g_\nu=1$ Werte η_ν mit $\eta_\nu\neq 0$ entsprechen. Sei $g_{\nu_1}=1$ und ν_1

minimal gewählt. Da g monoton wächst, so ist $g_\nu = 1$ für alle $\nu \geq \nu_1$, also gilt $\eta_\nu = 0$ für alle $\nu > \nu_1$ (dies folgt aus Satz 15 in § 4). Damit gilt $\eta_{\nu_1} = 1$ und $\eta_\nu = 0$ für $\nu \neq \nu_1$. Es folgt $\sum_0^\nu g_\mu \eta_\mu = 0$ für $\nu < \nu_1$ und $\sum_0^\nu g_\mu \eta_\mu = 1 = D$ für $\nu \geq \nu_1$. Aus $A_{\nu_0} = 0$ und $\eta_{\nu_0} = 0$ ergibt sich, daß $\nu_0 > \nu_1$ ist. Auf Grund dieser Ergebnisse ist nun $V_\nu(\eta) = -p + f_\nu(1+p)$ für $\nu < \nu_1$, $V_{\nu_1}(\eta) = -p + f_{\nu_1}$ und $V_\nu(\eta) = -p$ für $\nu > \nu_1$. Wegen $\xi_{\nu_0} \neq 0$ gilt $V_{\nu_0}(\eta) = v = -p$, denn es ist $\nu_0 > \nu_1$. Die Ungleichung $-p = v \geq V_{\nu_1}(\eta) = -p + f_{\nu_1}$ liefert dann $f_{\nu_1} = 0$, während $g_{\nu_1} = 1$ war. In dem gegebenen Duell auf dem Einheitsquadrat gibt es also ein x der Art, daß $f(x) = 0$ und $g(x) = 1$ ist. (Solche Spiele hatten wir schon auf S. 294 aus unseren Betrachtungen ausgeschlossen.) In diesem Spiel ist jede beliebige Strategie eine gute (also auch eine ε-gute) Strategie für den ersten Spieler. Irgendeine konvergente Folge von Strategien erfüllt daher den Satz 18 auf S. 304. Wir dürfen daher solche Spiele aus unseren Betrachtungen ausschließen. Dann gilt:

(2) Aus $\xi_\nu \neq 0$ und $\eta_\nu = 0$ folgt $A_\nu \neq 0$.

Es ist also entweder $\eta_\nu \neq 0$ oder $\eta_{\nu+1} \neq 0$. Wenn $\eta_\nu \neq 0$ ist, so folgt $U_\nu(\xi) = v \leq U_{\nu-1}(\xi)$, oder ausführlich:

$$C - g_\nu \left\{ p + C - (1+p) \sum_0^{\nu-1} f_\mu \xi_\mu \right\} + f_\nu g_\nu \xi_\nu$$

$$\leq C - g_{\nu-1} \left\{ p + C - (1+p) \sum_0^{\nu-1} f_\mu \xi_\mu \right\} - p f_{\nu-1} g_{\nu-1} \xi_{\nu-1},$$

oder

$$f_\nu g_\nu \xi_\nu + p f_{\nu-1} g_{\nu-1} \xi_{\nu-1} \leq (1+p)(g_\nu - g_{\nu-1}).$$

Berücksichtigen wir, daß $f_\nu g_\nu \geq \alpha > 0$ ist, so folgt

$$\xi_\nu \leq \frac{1+p}{\alpha}(g_\nu - g_{\nu-1}). \tag{3}$$

Wenn $\eta_{\nu+1} \neq 0$ ist, so folgt $U_{\nu+1}(\xi) = v \leq U_\nu(\xi)$, oder

$$C - g_{\nu+1} \left\{ p + C - (1+p) \sum_0^{\nu} f_\mu \xi_\mu \right\} + f_{\nu+1} g_{\nu+1} \xi_{\nu+1}$$

$$\leq C - g_\nu \left\{ p + C - (1+p) \sum_0^{\nu} f_\mu \xi_\mu \right\} - p f_\nu g_\nu \xi_\nu,$$

also

$$p f_\nu g_\nu \xi_\nu + f_{\nu+1} g_{\nu+1} \xi_{\nu+1} \leq (1+p)(g_{\nu+1} - g_\nu).$$

Also ist:

$$\xi_\nu \leq \frac{1+p}{p\alpha}(g_{\nu+1} - g_\nu). \tag{4}$$

Wir wollen (3) und (4) zusammenfassen zu

$$\xi_\nu \leq \frac{1+p}{\alpha}(g_\nu - g_{\nu-1}) + \frac{1+p}{p\alpha}(g_{\nu+1} - g_\nu). \tag{5}$$

Wir schätzen hier, wie überhaupt im weiteren Verlauf dieses Paragraphen, eher etwas zu reichlich ab, als daß wir uns auf zusätzliche Fallunterscheidungen einlassen.

Formel (5) gilt unter der Voraussetzung, daß $f_{\nu+1} > f_\nu$ ist.

2. Fall. $f_{\nu-1} < f_\nu = f_{\nu+1} = \cdots = f_{\nu+s-1} < f_{\nu+s}$. Wegen Satz 15 aus § 4 können wir annehmen, daß $\xi_{\nu+1} = \xi_{\nu+2} = \cdots = \xi_{\nu+s-1} = 0$ ist. Sei $\xi_\nu \neq 0$. Dann folgt:

$$V_\nu(\eta) = v \geq V_{\nu+s}(\eta),$$

oder

$$-pD + f_\nu\Big\{1 + pD - (1+p)\sum_0^\nu g_\mu \eta_\mu\Big\} + f_\nu g_\nu \eta_\nu$$

$$\geq -pD + f_{\nu+s}\Big\{1 + pD - (1+p)\sum_0^\nu g_\mu \eta_\mu\Big\} - (1+p) f_{\nu+s} \sum_{\nu+1}^{\nu+s-1} g_\mu \eta_\mu -$$

$$- p f_{\nu+s} g_{\nu+s} \eta_{\nu+s}.$$

Also ist:

$$f_\nu g_\nu \eta_\nu + p f_{\nu+s} g_{\nu+s} \eta_{\nu+s} + (1+p) f_{\nu+s} \sum_{\nu+1}^{\nu+s-1} g_\mu \eta_\mu \geq A_\nu (f_{\nu+s} - f_\nu).$$

Entweder ist $\eta_\nu \neq 0$ oder — auf Grund von (2) — wenigstens eines der $\eta_{\nu+l}$ mit $0 < l \leq s$ ungleich Null. Ist $\eta_\nu \neq 0$, so folgt die Ungleichung $U_\nu(\xi) = v \leq U_{\nu-1}(\xi)$, welche auf die Abschätzung $f_\nu g_\nu \xi_\nu \leq (1+p)(g_\nu - g_{\nu+1})$ führt. Ist $\eta_{\nu+l} \neq 0$ und $l \neq s$, so folgt $U_{\nu+l}(\xi) = v \leq U_\nu(\xi)$, d.h.

$$C - g_{\nu+l}\Big\{p + C - (1+p)\sum_0^\nu f_\mu \xi_\mu\Big\}$$

$$\leq C - g_\nu\Big\{p + C - (1+p)\sum_0^\nu f_\mu \xi_\mu\Big\} - p f_\nu g_\nu \xi_\nu,$$

oder

$$p f_\nu g_\nu \xi_\nu \leq (1+p)(g_{\nu+l} - g_\nu).$$

Ist $\eta_{\nu+s} \neq 0$ so ist $U_{\nu+s}(\xi) = v \leq U_\nu(\xi)$ und eine entsprechende Rechnung liefert die Abschätzung

$$f_{\nu+s} g_{\nu+s} \xi_{\nu+s} + p f_\nu g_\nu \xi_\nu \leq (1+p)(g_{\nu+s} - g_\nu).$$

Wir fassen diese Abschätzungen zusammen zu

$$\xi_\nu \leq \frac{1+p}{\alpha}(g_\nu - g_{\nu-1}) + \frac{1+p}{p\alpha}(g_{\nu+s} - g_\nu). \tag{6}$$

Hierbei war $f_{\nu-1} < f_\nu = f_{\nu+1} = \cdots = f_{\nu+s-1} < f_{\nu+s}$ vorausgesetzt.

3. Fall. Es bleibt noch die Möglichkeit zu betrachten, daß $f_{\nu-1} < f_\nu = f_{\nu+1} = \cdots = f_\nu$ und $x_\nu = 1$ ist. Wir wollen nun nachträglich festsetzen:

$$x_0 = \min\{x \mid f(x) = f(1)\}$$

soll zu jeder Teilung $\mathfrak{T}_n$ gezählt werden. Wir werden für ξ_ν keine Abschätzung brauchen, wenn x_ν ein Teilpunkt für jede Teilung $\mathfrak{T}_n$ ist. Der Fall 3 erledigt sich dann durch diese nachträgliche Festsetzung.

Auf Grund der Formeln (5) und (6) können wir nun zeigen, daß $\{F_n\}$ gleichgradig T-stetig ist. Gegeben sei ein $\delta > 0$. Man wähle ein N so, daß

$$\frac{1+p}{\alpha N} + \frac{2(1+p)}{p\alpha N} \leq \delta$$

ist. Wir wollen zeigen, daß $\mathfrak{T}_N$ eine δ-Teilung für alle F_n mit $n \geq N$ ist. Sei $x_k < x' \leq x'' < x_{k+1}$ und $n \geq N$. Dann folgt:

$$|F_n(x'') - F_n(x')| \leq \sum_{x' < x_\nu \leq x''} \xi_\nu \leq \sum_{x_k < x_\nu < x_{k+1}} \xi_\nu \leq \frac{1+p}{\alpha}\{g(x_{k+1}) - g(x_k)\} +$$

$$+ \frac{1+p}{p\alpha}\{g(x_{k+1}) - g(x_k)\} + \frac{1+p}{p\alpha} \cdot \frac{1}{N} < \frac{1+p}{\alpha N} + \frac{2(1+p)}{p\alpha N} \leq \delta.$$

Diese Abschätzung wollen wir noch erläutern. Der Term $\frac{1+p}{\alpha}\{g(x_{k+1}) - g(x_k)\}$ entsteht aus den ersten Termen in den Formeln (5) und (6). Der Term $\frac{1+p}{p\alpha}\{g(x_{k+1}) - g(x_k)\}$ entspricht den zweiten Termen in den Formeln (5) und (6) genau dann, wenn kein Konstanz-Intervall $[a; b)$ von $f(x)$ existiert der Art, daß $x_k < a < x_{k+1}$ und $b \geq x_{k+1}$ ist. Der Term $\frac{1+p}{p\alpha}\frac{1}{N}$ könnte dann fortfallen. Es bleibt also der Fall zu betrachten, daß solch ein Konstanz-Intervall existiert. Dabei kann nicht gelten, daß $b > x_{k+1}$ ist, denn dann müßte a ein Punkt der Teilung $\mathfrak{T}_N$ sein. [Wir hatten ja auf S. 299 festgesetzt: Wenn ein Punkt von $\mathfrak{T}$ in einem Konstanz-Intervall von $f(x)$ liegt, so soll auch der Anfangspunkt dieses Konstanz-Intervalles mit zur Teilung $\mathfrak{T}$ gehören.] Sei also $b = x_{k+1}$. Dann gilt

$$\sum_{x_k < x_\nu < x_{k+1}} \xi_\nu \leq \frac{1+p}{\alpha}\{g(x_{k+1}) - g(x_k)\} + \frac{1+p}{p\alpha}\{g(x_{k+2}) - g(x_k)\}$$

$$< \frac{1+p}{\alpha}\{g(x_{k+1}) - g(x_k)\} + \frac{1+p}{p\alpha}\{g(x_{k+1}) - g(x_k)\} + \frac{1+p}{p\alpha}\frac{1}{N},$$

denn es war vorausgesetzt, daß $g(x_{k+2}) - g(x_{k+1}) < 1/N$ ist. Damit ist die Gültigkeit der Abschätzung $|F_n(x'') - F_n(x')| < \delta$ gezeigt; $\mathfrak{T}_N$ ist eine δ-Teilung für alle F_n mit $n \geq N$.

Wir wollen nun zu $\mathfrak{T}_N$ alle diejenigen endlichvielen Punkte hinzunehmen, welche nötig sind, damit eine δ-Teilung für die endlichvielen F_n mit $n < N$ entsteht. Auf diese Weise kommen wir zu einer δ-Teilung für alle F_n. Die Folge $\{F_n\}$ ist also gleichgradig T-stetig.

Wir wollen die Ergebnisse dieses Paragraphen zusammenfassen und erhalten:

Satz 17. *Bei einem symmetrischen α-Duell gibt es eine Folge ε-guter Strategien, welche gegen eine gute Strategie konvergiert.*

Satz 18. *Bei einem α-Duell, welches sich aus einem Duell mit stetigem $g(y)$ und $p \neq 0$ herleitet, gibt es eine Folge ε-guter Strategien des ersten Spielers, welche gegen eine gute Strategie des ersten Spielers konvergiert.*

§ 6. Abrundung und Zusammenfassung der Ergebnisse

Wir wollen versuchen, die Ergebnisse des vorigen Paragraphen auf beliebige Duelle auszudehnen. Dazu beweisen wir den

Satz 19. *Gegeben sei ein Duell mit dem Kern*

$$K^*(x;y)=\varphi(x)-p\,\psi(y)\begin{cases}+p\,\varphi(x)\,\psi(y) & \text{für } x<y\\ +0 & \text{für } x=y\\ -\varphi(x)\,\psi(y) & \text{für } x>y,\end{cases}$$

wobei φ,ψ halbstetig nach oben, $0\leq\varphi,\psi\leq 1$ und $p\geq 0$ ist.

Sei $f(x)=\sup\limits_{\xi\in[0,x]}\varphi(\xi)$, $g(y)=\sup\limits_{\eta\in[0,x]}\psi(\eta)$ *und*

$$K(x;y)=f(x)-p\,g(y)\begin{cases}+p\,f(x)\,g(y) & \text{für } x<y\\ +0 & \text{für } x=y\\ -f(x)\,g(y) & \text{für } x>y.\end{cases}$$

Dann gilt: f und g sind monoton steigend, von rechts stetig und

$$\int K(x;y)\,dF(x)\leq\int K^*(x;y)\,dF(x)$$

und

$$\int K(x;y)\,dG(y)\geq\int K^*(x;y)\,dG(y)$$

für jede Strategie $F\in\mathfrak{C}_1$ und $G\in\mathfrak{C}_2$. Dabei sind die Klassen $\mathfrak{C}_1$ und $\mathfrak{C}_2$ in bezug auf das Duell mit dem Kern $K(x;y)$ gebildet.

Der Beweis erfolgt in 3 Schritten.

Schritt 1. Wir zeigen, daß f monoton steigend und von rechts stetig ist.

$$f(x+h)=\sup_{[0,x+h]}\varphi(\xi)=\max\left\{\sup_{[0,x]}\varphi(\xi);\ \sup_{(x,x+h]}\varphi(\xi)\right\}$$

$$=\max\left\{f(x);\ \sup_{(x,x+h]}\varphi(\xi)\right\}\geq f(x).$$

Es ist also $f(x+h)\geq f(x)$.

$$f(x+0)=\lim_{h\to 0}f(x+h)=\lim_{h\to 0}\max\left\{f(x);\ \sup_{(x,x+h]}\varphi(\xi)\right\}$$

$$=\max\left\{f(x);\lim_{h\to 0}\sup_{(x,x+h]}\varphi(\xi)\right\}=f(x),$$

denn

$$\lim_{h\to 0}\sup_{(x,x+h]}\varphi(\xi)\leq\lim_{h\to 0}\sup_{[x-h;x+h]}\varphi(\xi)=\lim_{\xi\to x}\sup\varphi(\xi)=\varphi(x)\leq f(x).$$

Es gilt also $f(x+0)=f(x)$.

Schritt 2. Wir wollen zeigen: Wenn $\varphi(x)\neq f(x)$ ist, so gibt es ein $\alpha>0$ der Art, daß für alle δ mit $0<\delta<\alpha$ die Gleichung $f(x-\delta)=f(x+\delta)$ gilt.

Sei $f(x)=a\neq\varphi(x)$, also $\varphi(x)<a$. Da $\varphi(x)$ halbstetig nach oben ist, so gibt es ein $\alpha>0$ so, daß für alle ξ mit $|\xi-x|<\alpha$ die Ungleichung $\varphi(\xi)<a$ erfüllt ist. (Man vgl. hierzu: Natanson [*10*] S. 410, Satz 3.) Nun können

wir wie folgt schließen. Wenn $0<\delta<\alpha$ und α wie oben gewählt wird, so gilt

$$f(x+\delta)=\max\left\{f(x);\ \sup_{(x,\,x+\delta]}\varphi(\xi)\right\}=f(x),$$

denn

$$\sup_{(x,\,x+\delta]}\varphi(\xi)\leq\sup_{[x,\,x+\delta]}\varphi(\xi)=\max_{[x,\,x+\delta]}\varphi(\xi)<a=f(x).$$

Ferner ist:

$$f(x)=\max\left\{f(x-\delta);\ \sup_{(x-\delta,\,x]}\varphi(\xi)\right\}=f(x-\delta),$$

denn

$$\sup_{(x-\delta,\,x]}\varphi(\xi)\leq\sup_{[x-\delta,\,x]}\varphi(\xi)=\max_{[x-\delta,\,x]}\varphi(\xi)<a=f(x).$$

Das heißt, die andere Möglichkeit, nämlich $f(x)=\sup\limits_{(x-\delta,\,x]}\varphi(\xi)$, führt auf einen Widerspruch. Damit ist gezeigt: Für $0<\delta<\alpha$ gilt $f(x-\delta)=f(x)=f(x+\delta)$.

Schritt 3. Sei $F\in\mathfrak{C}_1$. Nach der Definition von $\mathfrak{C}_1$ (S. 292) gilt dann $\int\limits_{(a,\,b)}dF=0$, falls $f(x)=c$ für $x\in[a;b\rangle$ ist. Wegen Schritt 2 können wir schließen: $\int\limits_{(a,\,b)}dF=0$, wenn $f(x)\neq\varphi(x)$ für $x\in(a;b)$ gilt. Wir wollen jetzt die Ungleichung $\int K^*(x;y)\,dF(x)\geq\int K(x;y)\,dF(x)$ beweisen.

$$\begin{aligned}\int K^*(x;y)\,dF(x)-\int K(x;y)\,dF(x)&=\int\varphi\,dF-\psi(y)\{p+\int\varphi\,dF-(1+p)\psi(y)\times\\&\times\int^{y}\varphi\,dF\}+\varphi(y)\,\psi(y)\,\sigma_F(y)-\int f\,dF+g(y)\{p+\int f\,dF-(1+p)\,g(y)\times\\&\times\int^{y}f\,dF\}-f(y)\,g(y)\,\sigma_F(y)=\int(\varphi-f)\,dF+\\&+\{g(y)-\psi(y)\}\cdot\{p+\int f\,dF-(1+p)\cdot\int^{y}f\,dF-f(y)\,\sigma_F(y)\}+\\&+\psi(y)\{\int(f-\varphi)\,dF-(1+p)\cdot\int^{y}(f-\varphi)\,dF-\big(f(y)-\varphi(y)\big)\,\sigma_F(y)\}.\end{aligned}$$

Dieser Ausdruck ist größer oder gleich Null, denn es ist:

$$g(y)-\psi(y)\geq 0$$

und

$$\begin{aligned}p+\int f\,dF-(1+p)\cdot\int^{y}f\,dF-f(y)\,\sigma_F(y)&\geq(1+p)\{\int f\,dF-\\&-\int^{y}f\,dF-f(y)\,\sigma_F(y)\}\geq 0.\end{aligned}$$

Schließlich gilt $\int\limits^{y}(f-\varphi)\,dF=0$ und $\{f(y)-\varphi(y)\}\,\sigma_F(y)=0$, denn entweder ist $f(y)=\varphi(y)$, oder y befindet sich im Inneren eines Intervalles vom F-Maß Null. Es ist also:

$$\int K(x;y)\,dF(x)\leq\int K^*(x;y)\,dF(x).$$

Genau so folgt:

$$\int K(x;y)\,dG(y)\geq\int K^*(x;y)\,dG(y).$$

Damit ist Satz 19 bewiesen.

Satz 19 sagt aus: Wenn ein beliebiges Duell gegeben ist, so läßt sich stets ein Duell mit monoton steigenden, von rechts stetigen Treffwahrscheinlich-

keiten angeben, so daß jede für das monotone Duell gute oder ε-gute Strategie (aus den Klassen $\mathfrak{C}_1$ bzw. $\mathfrak{C}_2$) auch gut oder ε-gut für das gegebene Duell ist. Aussagen über monotone Duelle lassen sich also auf allgemeine Duelle übertragen. Dann folgt aus Satz 14 in § 4, S. 299

Satz 20. *Jedes Duell hat einen Wert.*

Wir betrachten jetzt Duelle mit monotonen Treffwahrscheinlichkeiten und mit $p \neq 0$. Der Satz 11 aus § 3 (S. 295) sagt aus, daß wir uns bei der Betrachtung von guten oder ε-guten Strategien eines monotonen Duells mit $p \neq 0$ auf die Klasse der α-Duelle beschränken können, denn bei den guten und ε-guten Strategien dürfen wir nach § 3 stets annehmen, daß sie zu den Klassen $\mathfrak{C}_1$ bzw. $\mathfrak{C}_2$ gehören. Da auch die speziellen ε-guten Strategien des ersten Spielers, welche wir in § 5 betrachtet haben, zur Klasse $\mathfrak{C}_1$ gehörten, so übertragen sich die Ergebnisse des § 5 auf alle Duelle mit $p \neq 0$ und monotonen Treffwahrscheinlichkeiten, und damit nach dem Satz 19 dieses Paragraphen auf alle Duelle mit $p \neq 0$. Wir haben damit folgendes Ergebnis erhalten:

Satz 21. *Bei symmetrischen Duellen hat jeder Spieler eine gute Strategie. Bei Duellen mit $p \neq 0$ und stetigem g hat der erste Spieler eine gute Strategie. Bei beiden Duelltypen gibt es eine Folge ε-guter Strategien, welche gegen eine gute Strategie konvergiert.*

Anmerkung. Natürlich gilt genau so: Bei Duellen mit $p \neq 0$ und stetigem f hat der zweite Spieler eine gute Strategie.

Die Existenz der guten Strategien des ersten Spielers konnte nicht bewiesen werden ohne die Voraussetzung, daß g stetig sei. Ein solcher Beweis ist auch ohne Zusatzvoraussetzungen nicht möglich. Wir können nämlich leicht ein Duell mit unstetigem g angeben, welches keine guten Strategien für den ersten Spieler besitzt. Wir geben ein Beispiel: Sei

$$f(x) = x \quad \text{und} \quad g(y) = \begin{cases} \frac{1}{2} & \text{für } y < \frac{1}{2} \\ 1 & \text{für } y \geq \frac{1}{2} \end{cases}$$

und $p = 1$. Wählt der zweite Spieler die Strategie $G(y) = \frac{4}{5} I_0(y) + \frac{1}{5} I_{\frac{1}{2}}(y)$, so ist

$$V(x; G(y)) = \begin{cases} -\frac{3}{5} + (\frac{4}{5}) \cdot x & \text{für } 0 \leq x < \frac{1}{2} \\ -\frac{9}{10} & \text{für } x = \frac{1}{2} \\ -\frac{3}{5} + (\frac{2}{5}) \cdot x & \text{für } \frac{1}{2} < x \leq 1. \end{cases}$$

Man hat also: $V(x; G(y)) \leq -\frac{1}{5}$. Wählt der erste Spieler die Strategie $F_\varepsilon(x) = \frac{4}{5} I_{\frac{1}{2}-\varepsilon}(x) + \frac{1}{5} I_1(x)$, so ist

$$U(F_\varepsilon(x), y) = \begin{cases} -\frac{1}{5} + (\frac{2}{5})\,\varepsilon & \text{für } 0 \leq y < \frac{1}{2} - \varepsilon \\ -(\frac{4}{5})\,\varepsilon & \text{für } y = \frac{1}{2} - \varepsilon \\ -\frac{1}{5} - (\frac{6}{5})\,\varepsilon & \text{für } \frac{1}{2} - \varepsilon < y < \frac{1}{2} \\ -\frac{1}{5} - (\frac{8}{5})\,\varepsilon & \text{für } \frac{1}{2} \leq y < 1 \\ -(\frac{8}{5})\,\varepsilon & \text{für } y = 1. \end{cases}$$

Man hat also: $U(F_\varepsilon(x); y) \geq -\frac{1}{5} - (\frac{8}{5})\,\varepsilon$.

Daraus entnehmen wir: Der Wert des Spieles ist $v=-\frac{1}{3}$, $G(y)$ ist eine gute Strategie für den zweiten Spieler, und $F_\varepsilon(x)$ ist eine ε-gute Strategie für den ersten Spieler.

Wir können nun notwendige Bedingungen für eine gute Strategie des ersten Spielers aufstellen. Eine gute Strategie $F(x)$ darf nur diejenigen Punkte des Intervalles $[0;1]$ mit Gewicht belegen, an welchen $V(x;G(y))=v$ ist, also nur den Punkt $x=1$. Als gute Strategie für den ersten Spieler käme daher nur $F(x)=I_1(x)$ in Frage, wir rechnen aber leicht nach, daß

$$U(I_1(x);\tfrac{1}{2})=-1<-\tfrac{1}{3}=v$$

ist. Es gibt also keine gute Strategie für den ersten Spieler.

Wir können das Beispiel abändern, indem wir setzen:

$$f(x)=\begin{cases}\alpha & \text{für } x\leq\alpha\\ x & \text{für } x\geq\alpha.\end{cases}$$

Solange $\alpha<\frac{1}{2}$ ist, hat auch jetzt der erste Spieler keine gute Strategie. Für $\alpha=\frac{1}{2}$ dagegen gibt es eine gute Strategie für den ersten Spieler. Ein Kriterium dafür, wann bei einem Duell mit unstetigem g der erste Spieler eine gute Strategie besitzt, ist noch nicht bekannt.

Literatur

[1] BLACKWELL-GIRSHICK: Theorie of games and statistical decisions. New York 1954. — [2] BURGER, E.: Spieltheoretische Behandlung eines Reklameproblems. Mitteilungsblatt math. Statistik **6**, 39—52 (1954). — [3] CRAMÉR, H.: Mathematical methods of statistics. Princeton 1946. — [4] DANSKIN-GILLMAN: A game over function-space. Riv. di mat. Univ. Parma **4**, 83—94 (1953). — [5] FRÉCHET, M.: Recherches théoriques modernes, premier livre. (Tome I, Fasc. II von BOREL: Traité du calcul des probabilités.) Paris 1937. — [6] GILLMAN, L.: Operations analysis and the theory of games: An advertising exemple. J. Amer. Statist. Assoc. **45**, 514—545 (1950). — [7] GNEDENKO-KOLMOGOROV: Limit distributions for sums of independent random variables. Cambridge Mass. 1954. — [8] KARLIN, S.: Reduction of certain classes of games to integral equations. Contributions to the theory of games II. Ann. of Math. Stud. **28**, 125—158 (1953). — [9] MCKINSEY, J. C. C.: Introduction to the theory of games. New York 1952. — [10] NATANSON, I. P.: Theorie der Funktionen einer reellen Veränderlichen. Berlin 1954. — [11] NEUMANN-MORGENSTERN, v.: Theory of games and economic behavior, 2. Aufl. Princeton 1947. — [12] SHIFFMAN, M.: Games of timing. Contributions to the theory of games II. Ann. of Math. Stud. **28**, 97—124 (1953).

Tübingen, Mathematisches Institut der Universität

(Eingegangen am 3. November 1955)

DOLPH, C. L., and R. K. RITT
Math. Zeitschr. Bd. 65, S. 309—326 (1956)

The Schwinger variational principles for one-dimensional quantum scattering

Dedicated to the Memory of HERMANN WEYL

By

C. L. DOLPH and R. K. RITT

§ 1. Introduction

The success of variational methods in the solution of scattering problems [9] has aroused an interest in the rigorous formulation and deeper understanding of these methods [1], [4], [6]. In this paper, we examine the variational principle for the one-dimensional quantum scattering of a plane wave with positive energy, k^2, by a positive potential, $V(x)$, for which $\int_{-\infty}^{\infty} V(x)\,dx < \infty$. We find:

1. If $\int_{-\infty}^{\infty} V(x)\,dx < 2k$, the Schwinger variational principle is valid; i.e., the behavior of the transmitted wave for a large x can be determined by the stationary value of an appropriate bilinear integral form, subject to a normalizing constraint. If $\int_{-\infty}^{\infty} V(x)\,dx \geq 2k$ an additional hypothesis must be imposed before the validity of the principle can be established; namely, the potential must not be "perfectly transparent" to the plane wave. To the best of our knowledge, this condition does not appear explicitly in the literature, although it is tacitly present in all the discussions of the principle. We call attention to an elementary example to show that a perfectly transparent potential can occur.

2. It is known [7] that in the one-dimensional case both the transmission and reflection coefficients can be characterized *formally* as stationary points of a variational principle. In the case of the reflection coefficient, the variational principle is symmetric with respect to a real inner product or, since it is convenient to use the theory of complex Hilbert space, quadratic with respect to a Hermitian inner product. In the case of the transmission coefficient, the variational principle is bilinear with respect to either inner product. However, as we will demonstrate, it can be reformulated as a quadratic variational principle in a suitably defined product Hilbert space. This reformulation not only makes a geometrical interpretation possible but in addition it results in unifying the theory since, when it is considered abstractly, it includes the variational principle for the reflection amplitude as a special case.

It should be noticed that our formulation differs from the usual one, in which the competing functions do not lie in a Hilbert space since they do not vanish at infinity. However, by multiplying the integral equation by $U(x) = [V(x)]^{\frac{1}{2}}$ if $V(x) \geq 0$, and by $U_1(x) = [|iV(x)|]^{\frac{1}{2}}$ if $V(x) \leq 0$, this formulation can be restated in terms of a Hilbert space. The first of these devices has been used by Kato [5]. Throughout this paper we will assume $V(x) \geq 0$ and work with $U(x)$. It is a simple matter for the reader to verify that the results and method of proof are similar under the alternate hypothesis.

3. In the case of the transmission coefficient, the stationary points form a one (complex) parameter curve on the constraining surface. These points can be divided into two classes, according to the behavior of the imaginary part of the quadratic form. At points of the first class, the imaginary part exhibits a local saddle point behavior; i.e., we express this imaginary part as a real quadratic form on the direct sum of the two real Hilbert spaces, and show that its projections onto these two spaces have, respectively, a local minimum and a local maximum. At points of the second class, one of these two projections has a global maximum and nothing can be said about the other.

In the case of the reflection coefficient, the situation is much simpler geometrically. There is only one class of possible stationary points and it forms a one-dimensional linear manifold. Here the projections of the imaginary part of the corresponding real quadratic form on the direct sum of the two corresponding real Hilbert spaces always leads to a global minimum and a global maximum so that there is always a global saddle point behavior. On the other hand, in this case we obtain less information concerning the reflection coefficient itself, since there appears to be no useful analog of the cross-section theorem from which to deduce its properties.

4. For the sake of clarity and at the expense of some unity in the treatment we first discuss the more difficult case of the transmission coefficient, reserving the treatment of the reflection coefficient for the last section of the paper. The outline of the paper is as follows: § 2 is devoted to a discussion of the pertinent properties of the associated Fredholm integral equations which underlie our entire theory; § 3 is devoted to a discussion to the relation of these equations with the scattering problem and to an example of a "transparent" potential; § 4 discusses the bilinear variational principle for the transmission coefficient in Hilbert space; § 5 discusses the quadratic formulation of the same principle in a product Hilbert space; § 6 is devoted to the characterization of the stationary points for the transmission coefficient; and § 7 is given over to the corresponding formulation and discussion appropriate for the reflection coefficient.

The notation will, for the most part, be self-explanatory. The symbol $(*, *)$ will always be used to denote the inner product, regardless of the space with which we are dealing; the objects which appear in place of the asterisks will identify the space. We freely use the δ of the calculus of variations;

this is equivalent to using Gateaux differentials [3]. When we say that a function is stationary we mean that its Gateaux differential is zero. The reader may be troubled by the extensive use of conjugations: in some cases the ordinary conjugation in a complex function space; in others a more abstract conjugation [8]. This could be avoided by using a symmetric, rather than a Hermitian symmetric, inner product, but this would involve a re-orientation for the reader familiar with Hilbert spaces; we have chosen what appears to be the lesser of two evils.

The argument does not depend upon the existence of a biorthogonal expansion. It can be proved, using methods of Dunford, Schwartz, Werner, and Bade [2], that the basic integral operator does have a biorthogonal spectral resolution. Thus, there is a possibility that the problem could be formulated analogously to the Rayleigh-Ritz procedure [1].

§ 2. The Fredholm equations

We shall consider the operator on $L^{(2)}(-\infty, \infty)$ defined by

$$K\varphi = \frac{1}{2ki} \int_{-\infty}^{\infty} U(x)\, e^{ik|x-y|}\, U(y)\, \varphi(y)\, dy \tag{2.1}$$

where $U(x)$ is a real valued function in $L^{(2)}(-\infty, \infty)$.

LEMMA 1. *K is a bounded linear operator. The operator norm of K, $|||K|||$, satisfies*

$$\|K\| \leq \frac{1}{2k} \|U(x)\|^2. \tag{2.2}$$

K is compact. The adjoint of K, K^* satisfies

$$K^*\varphi = \overline{K\bar{\varphi}}. \tag{2.3}$$

If φ_0 is defined by

$$\varphi_0(x) = U(x)\, e^{ikx}, \tag{2.4}$$

then

$$(K\varphi, \psi) - (\varphi, K\psi) = \frac{1}{2ki} \{(\varphi, \varphi_0)(\varphi_0, \psi) + (\varphi, \bar{\varphi}_0)(\bar{\varphi}_0, \psi)\}, \tag{2.5}$$

for all φ and ψ in $L^{(2)}(-\infty, \infty)$.

PROOF. (2.2) is obtained from Schwartz's inequality. K is generated by a Hilbert-Schmidt kernel and is therefore compact. (2.3) is obtained from inspection. To prove (2.5) we integrate by parts as follows:

$$\begin{aligned}(K\varphi, \psi) &= \frac{1}{2ki} \int_{-\infty}^{\infty} \overline{\psi(x)} \int_{-\infty}^{\infty} U(x)\, e^{ik|x-y|} U(y)\, \varphi(y)\, dy\, dx \\ &= \frac{1}{2ki} \int_{-\infty}^{\infty} U(x)\, \overline{\psi(x)} \left\{ e^{ikx} \int_{-\infty}^{x} e^{-iky}\, U(y)\, \varphi(y)\, dy + \right. \\ &\quad \left. + e^{-ikx} \int_{x}^{\infty} e^{iky}\, U(y)\, \varphi(y)\, dy \right\} dx\end{aligned}$$

$$= \frac{1}{2ki}\left[\int_{-\infty}^{x} U(y)\,\overline{\psi(y)}\,e^{iky}\,dy \int_{-\infty}^{x} e^{-iky}\,U(y)\,\varphi(y)\,dy\right]_{-\infty}^{\infty} -$$

$$- \frac{1}{2ki}\left[\int_{x}^{\infty} U(y)\,\overline{\psi(y)}\,e^{-iky}\,dy \int_{x}^{\infty} e^{iky}\,U(y)\,\varphi(y)\,dy\right]_{-\infty}^{\infty} -$$

$$- \frac{1}{2ki}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{x} U(y)\,\overline{\psi(y)}\,e^{iky}\,dy\right)e^{-ikx}\,U(x)\,\varphi(x)\,dx -$$

$$- \frac{1}{2ki}\int_{-\infty}^{\infty}\left(\int_{x}^{\infty} U(y)\,\overline{\psi(y)}\,e^{-iky}\,dy\right)e^{ikx}\,U(x)\,\varphi(x)\,dx$$

$$= \frac{1}{2ki}\{(\varphi,\varphi_0)(\varphi_0,\psi) + (\varphi,\varphi_0)(\varphi_0,\psi)\} +$$

$$+ \int_{-\infty}^{\infty}\varphi(x)\left(\overline{\frac{1}{2ki}\int_{-\infty}^{\infty} U(x)\,e^{-ik|x-y|}\,U(y)\,\psi(y)\,dy}\right)dx$$

$$= \frac{1}{2ki}\{(\varphi,\varphi_0)(\varphi_0,\psi) + (\varphi,\bar{\varphi}_0)(\bar{\varphi}_0,\psi)\} + (\varphi, K\psi).$$

LEMMA 2. *The imaginary part of every eigenvalue of K is non-positive, and the eigenfunction φ corresponds to a real eigenvalue if and only if $(\varphi,\varphi_0) = (\varphi,\bar{\varphi}_0) = 0$.*

PROOF. If $K\varphi = \lambda\varphi$, then from (2.5),

$$(\lambda - \bar{\lambda})\,\|\varphi\|^2 = \frac{1}{2ki}\{|(\varphi,\varphi_0)|^2 + |(\varphi,\bar{\varphi}_0)|^2\}.$$

Now, we define the operators

$$F = \frac{K - K^*}{2i}, \qquad G = \frac{K + K^*}{2}. \tag{2.6}$$

LEMMA 3. *F and G are bounded, self-adjoint, compact linear operators. $K = G + iF$. F is negative semi-definite. The range of F is the two dimensional space, H_0, spanned by φ_0 and $\bar{\varphi}_0$.*

PROOF. The first two statements are trivial. Using (2.5) and (2.6) we have

$$F\varphi = -\frac{1}{4k}\{(\varphi,\varphi_0)\,\varphi_0 + (\varphi,\bar{\varphi}_0)\,\bar{\varphi}_0\}, \tag{2.7}$$

which proves the lemma.

Now, suppose $K\varphi = \varphi$. Then, from lemma 2 and (2.7), $F\varphi = 0$. Therefore, $G\varphi = K\varphi = \varphi$. Conversely, if $G\varphi = \varphi$ and $F\varphi = 0$, $K\varphi = \varphi$.

LEMMA 4. *Let H_1 be the intersection of the null space of $I - G$ (I = identity) and $H_0^{\perp}$. Then H_1 is the null space of $A = I - K$, and $H = H_1^{\perp}$ remains invariant under K. H_1 is also the null space for $A^* = I - K^*$, and H remains invariant under K^*.*

PROOF. The paragraph which precedes the lemma shows that $A\varphi = 0$ if and only if $F\varphi = 0$ and $G\varphi = \varphi$; this is equivalent to $F^*\varphi = 0$ and $G^*\varphi = \varphi$;

thus, H_1 is also the null space for A^*. Let ψ be in H. Then $(\psi, \varphi) = 0$ for all φ in H_1. But $(K\psi, \varphi) = (\psi, K^*\varphi) = (\psi, G^*\varphi) = 0$, for all φ in H_1. Therefore $K\psi$ is in H. Similarly, $K^*\psi$ is in H. This completes the proof.

THEOREM 1. *The operators $A = I - K$, $A^* = I - K^*$, when regarded as operators on H both have bounded inverses. In particular, the Fredholm equations*

$$A\varphi = \varphi_0; \quad A^*\bar{\psi} = \varphi_0, \tag{2.8}$$

have unique solutions in H.

PROOF. This follows directly from lemma 4. It should be observed that if the inhomogeneous terms in (2.8) are replaced by any functions in H, the theorem continues to be true. We use $\bar{\psi}$ rather than ψ in (2.8) for later convenience.

THEOREM 2. *If φ_1 and $\bar{\psi}_1$ are the solutions of* (2.8) *then*

$$(\varphi_1, \varphi_0) = (\varphi_0, \bar{\psi}_1) \quad \textit{(Reciprocity)}. \tag{2.9}$$

If $\|U(x)\|^2 < 2k$, then $(\varphi_1, \varphi_0) \neq 0$.

PROOF. $(\varphi_1, \varphi_0) = (\varphi_1, A^*\bar{\psi}_1) = (A\varphi_1, \bar{\psi}_1) = (\varphi_0, \bar{\psi}_1)$; if $A\varphi_1 = \varphi_1 - K\varphi_1 = \varphi_0$, and $(\varphi_1, \varphi_0) = 0$, then $\|\varphi_1\|^2 = (K\varphi_1, \varphi_1) \leq |||K||| \, \|\varphi_1\|^2$. Therefore $|||K||| \geq 1$. But if $\|U(x)\|^2 < 2k$, this contradicts (2.2). The theorem is proved.

In what follows, regardless of the value of $\|U(x)\|$, we shall assume that φ_1 satisfies

$$(\varphi_1, \varphi_0) \neq 0. \tag{2.10}$$

As we shall see in § 3 this is equivalent to assuming that the potential is not perfectly transparent. Restricting ourselves to H rather than studying the equations (2.8) on all of $L^{(2)}(-\infty, \infty)$ is not essential, as we shall see in Theorem 4 below. We do so as a matter of elegance, since otherwise we would have to keep in mind the finite dimensional (G is compact) subspace, H_1, and phrase our statements accordingly.

§ 3. Discussion of the preceding formulations

We shall now connect the Fredholm equations (2.8) with the problem of one-dimensional quantum scattering. The fundamental analytic object one has to deal with is the time independent Schrödinger equation

$$\omega'' + [k^2 - V(x)]\,\omega = 0, \tag{3.1}$$

subject to the conditions

$$\lim_{x \to \infty} [\omega' - ik\omega] = 0, \tag{3.2}$$

$$\lim_{x \to -\infty} [\omega' + ik\omega]\, e^{-ikx} = 2ik. \tag{3.3}$$

Solving this problem is equivalent to solving the integral equation

$$\omega(x) = e^{ikx} + \frac{1}{2ki}\int_{-\infty}^{\infty} e^{ik|x-y|}\, V(y)\,\omega(y)\,dy. \tag{3.4}$$

If we assume that $\int_{-\infty}^{\infty} |V(y)|\,dy < \infty$, this is a Fredholm equation of the second kind in the space of essentially bounded functions. Now, if $V(y) \geq 0$, we avoid the difficulty of dealing with a space which is not self-dual, by setting $U(x) = [V(x)]^{\frac{1}{2}}$, $\varphi(x) = U(x)\,\omega(x)$, $\varphi_0(x) = U(x)\,e^{ikx}$, and multiplying (3.4) by $U(x)$. This gives the first equation in (2.8). It is clear that the solution, φ_1, of (2.8) determines the solution of (3.4), except on the set for which $U(x) = 0$. On this set $\omega(x)$ can be determined from (3.1) and the usual requirements of continuity of $\omega(x)$ and $\omega'(x)$.

The critical quantity we wish eventually to determine is

$$\frac{1}{\lambda_1} = \int_{-\infty}^{\infty} \omega(x)\, V(x)\, e^{-ikx}\,dx, \tag{3.5}$$

where $\omega(x)$ is the solution of (3.4). $1 + \frac{1}{2ki\lambda_1}$ is the complex amplitude of the transmitted radiation. But $\frac{1}{\lambda_1}$ is exactly the same as (φ_1, φ_0). Thus we lose nothing by dealing with (2.8) rather than (3.4).

Now what is the structure of the space H_1? To understand this we shall interpret H_1 from the point of view of equation (3.1). Suppose that $\varphi(x)$ simultaneously satisfies

$$\varphi = K\varphi, \tag{3.6}$$

$$(\varphi, \varphi_0) = (\varphi, \bar{\varphi}_0) = 0. \tag{3.7}$$

Then the corresponding $\omega(x)$ satisfies

$$\omega(x) = \frac{1}{2ki}\int_{-\infty}^{\infty} e^{ik|x-y|}\, V(y)\,\omega(y)\,dy, \tag{3.8}$$

$$\int_{-\infty}^{\infty} \omega(x)\, V(x)\, e^{ikx}\,dx = \int_{-\infty}^{\infty} \omega(x)\, V(x)\, e^{-ikx}\,dx = 0. \tag{3.9}$$

But then we can write

$$\omega(x) = \frac{1}{2ki}\left\{e^{ikx}\int_{-\infty}^{x} e^{-iky}\, V(y)\,\omega(y)\,dy + e^{-ikx}\int_{-x}^{\infty} e^{iky}\, V(y)\,\omega(y)\,dy\right\}, \tag{3.10}$$

$$\omega'(x) = \tfrac{1}{2}\left\{e^{ikx}\int_{-\infty}^{x} e^{-iky}\, V(y)\,\omega(y)\,dy - e^{-ikx}\int_{-x}^{\infty} e^{iky}\, V(y)\,\omega(y)\,dy\right\}, \tag{3.11}$$

$$\omega'' + [k^2 - V(x)]\,\omega = 0. \tag{3.12}$$

Combining (3.10) and (3.11), and using (3.9) we secure the following conditions which must be satisfied by $\omega(x)$:

$$\lim_{x\to\infty}[ik\omega+\omega']=0; \qquad \lim_{x\to\infty}[ik\omega-\omega']=0, \tag{3.13}$$

$$\lim_{x\to-\infty}[ik\omega+\omega']=0; \qquad \lim_{x\to-\infty}[ik\omega-\omega']=0. \tag{3.14}$$

From (3.13) and (3.14) it is apparent that $\omega(x)$ and $\omega'(x)$ must approach zero as x approaches both $+\infty$ and $-\infty$. The $\omega(x)$ then correspond to certain *bound states* of the system; i.e., solutions of the Schrödinger equation which together with their derivatives vanish at $+\infty$ and $-\infty$. In limiting ourselves to H we exclude from consideration those $\varphi(x)$ which correspond to these bound states. It is only when we do this that we guarantee the uniqueness of the solution of (2.8). There may well exist other bound states. These will correspond to those eigenfunctions of K which have real eigenvalues. This is easily verified by making use of lemma 2 and repeating the above argument. It serves no useful purpose to exclude these other bound states.

The condition (2.10) is, as we shall see, forced upon us if we wish to find (3.5) using a variational principle. In theorem 2 we proved that if the potential is "sufficiently weak", i.e.,

$$\int_{-\infty}^{\infty} V(x)\,dx < 2k, \tag{3.15}$$

then (2.10) is automatically satisfied. But what can be said if (3.15) fails to hold? Let us suppose that $(\varphi_1, \varphi_0)=0$. Since $\varphi_1=\varphi_0+K\varphi_1$, $(K\varphi_1, \varphi_1)=(\varphi_1, \varphi_1)=(\varphi_1, K\varphi_1)$. Consequently, using (2.5), $(\varphi_1, \bar{\varphi}_0)=0$. Then the corresponding $\omega(x)$, which is a solution of (3.4), satisfies the conditions (3.9). As in the discussion which led to (3.13) and (3.14), we obtain

$$\lim_{x\to\infty}[ik\omega+\omega']e^{-ikx}=\lim_{x\to-\infty}[ik\omega+\omega']e^{-ikx}=2ik, \tag{3.16}$$

$$\lim_{x\to\infty}[ik\omega-\omega']=\lim_{x\to-\infty}[ik\omega-\omega']=0. \tag{3.17}$$

Therefore, we have

$$\lim_{x\to\infty}\omega e^{-ikx}=\lim_{x\to-\infty}\omega e^{-ikx}=1. \tag{3.18}$$

It is (3.18) which supplies a precise meaning to the assertion that $V(x)$ is "perfectly transparent" to the incoming plane wave e^{ikx}. For large positive values of x, no scattering will be detected, and for large negative values of x, no reflection will be detected. That this can happen, even in a simple type of situation, can be seen by letting $k=4\pi$, and defining

$$V(x)=\begin{cases}12\pi^2 & 0\le x\le 1\\ 0 & \text{otherwise}\end{cases}.$$

then the function

$$\omega(x) = \begin{cases} \frac{3}{2} e^{2\pi i x} - \frac{1}{2} e^{-2\pi i x} & 0 \leq x \leq 1 \\ e^{4\pi i x} & \text{otherwise} \end{cases}$$

is the solution of (3.4). In this case, the variational principle will not apply, as will become apparent in § 4. It is an open question to find the general conditions on $V(x)$ which will guarantee (2.10). Trivially, (2.10) fails to be satisfied if and only if $(A^{-1}\varphi_0, \varphi_0) = 0$, but it is not obvious how to extract from this a sharp criterion. For a square wave potential, the question is easy to settle, but we shall not trouble the reader with the details.

§ 4. The variational principle for the transmission coefficient

We shall consider the variational technique of finding (φ_1, φ_0). Consider the bilinear form

$$J = (A\varphi, \overline{\psi}). \tag{4.1}$$

Let us attempt to make J stationary, subject to the normalizing constraint

$$(\varphi, \varphi_0)(\varphi_0, \overline{\psi}) = 1. \tag{4.2}$$

To do this, we introduce the Lagrange multiplier, λ, and set

$$J_\lambda = J - \lambda(\varphi, \varphi_0)(\varphi_0, \overline{\psi}). \tag{4.3}$$

Then the first variation of J_λ is given by

$$\begin{cases} \delta J_\lambda = (\delta\varphi, A^*\overline{\psi}) + (A\varphi, \delta\overline{\psi}) - \lambda(\delta\varphi, \varphi_0)(\varphi_0, \overline{\psi}) - \lambda(\varphi, \varphi_0)(\varphi_0, \delta\overline{\psi}) \\ \quad = (\delta\varphi, A^*\overline{\psi} - \overline{\lambda}(\overline{\psi}, \varphi_0)\varphi_0) + (A\varphi - \lambda(\varphi, \varphi_0)\varphi_0, \delta\overline{\psi}). \end{cases} \tag{4.4}$$

Therefore J_λ is stationary if and only if

$$A\varphi = \lambda(\varphi, \varphi_0)\varphi_0, \qquad A^*\overline{\psi} = \overline{\lambda}(\overline{\psi}, \varphi_0)\varphi_0. \tag{4.5}$$

THEOREM 3 (Schwinger's variational principle). *There exists exactly one value of λ, λ_1, such that J_λ can be made stationary. λ_1 is equal to the value of J when J is made stationary subject to the constraint* (4.2). *Simultaneously, $\lambda_1 = \frac{1}{(\varphi_1, \varphi_0)}$. The pair, φ, ψ make J stationary subject to* (4.2) *if and only if there exists an $\alpha \neq 0$ such that $\varphi = \alpha\lambda_1\varphi_1$, $\psi = \frac{1}{\alpha}\lambda_1\psi_1$.*

PROOF. Since we are assuming (2.10), let $\lambda = \frac{1}{(\varphi_1, \varphi_0)} = \frac{1}{(\varphi_0, \overline{\psi}_1)}$. The last equality follows from theorem 2. Then the pair φ_1, ψ_1 satisfy (4.5). Thus, J_λ can be made stationary. On the other hand, if J_λ can be made stationary, there exists a pair, φ, ψ such that (4.5) holds. From theorem 1, A and A^* have bounded inverses, so $\lambda(\varphi, \varphi_0) \neq 0$ and $\overline{\lambda}(\overline{\psi}, \varphi_0) \neq 0$. Then it is an easy matter, from the uniqueness of the solutions of (2.8), to see that $\varphi = \lambda(\varphi, \varphi_0)\varphi_1$, $\psi = \lambda(\varphi_0, \overline{\psi})\psi_1$. But then taking the inner product of the first of these with

φ_0, we find that $\lambda(\varphi_1, \varphi_0) = 1$. Thus $\lambda = \lambda_1$ is uniquely determined. If φ and ψ satisfy (4.2), from (4.5) we obtain $\lambda = (A\varphi, \bar{\psi})$, and from above, we obtain $\varphi = \alpha\lambda_1\varphi_1$, $\psi = \frac{1}{a}\lambda_1\psi_1$, where $\alpha = (\varphi, \varphi_0)$. Finally, for every $\alpha \neq 0$, $\alpha\lambda_1\varphi_1$, $\frac{1}{a}\lambda_1\psi_1$ satisfy both (4.5) and (4.2). To emphasize the importance of (2.10), we state explicitly:

COROLLARY. *If $(\varphi_1, \varphi_0) = 0$ there does not exist a λ such that J_λ can be made stationary, and consequently, J cannot be made stationary subject to* (4.2).

The next theorem shows that in the application of the variational principle it is not necessary to distinguish between H and $L^{(2)}(-\infty, \infty)$. Our use of the space H has been a convenience for the easier statement and proof of theorem 3.

THEOREM 4. *Let φ and $\bar{\psi}$ be in $L^{(2)}(-\infty, \infty)$. Then φ and $\bar{\psi}$ have unique representations as $\varphi' + \varphi''$, $\bar{\psi}' + \bar{\psi}''$, respectively, where φ' and $\bar{\psi}'$ are in H_1, $A\varphi''$ and $\bar{\psi}''$ are in H $(A\varphi, \bar{\psi}) = (A\varphi'', \bar{\psi}'')$, and $(\varphi, \varphi_0)(\varphi_0, \bar{\psi}) = (\varphi'', \varphi_0)(\varphi_0, \bar{\psi}'')$.*

PROOF. The proof follows easily from lemma 3. In the language of § 3, in the application of the variational principle, the failure to factor out those functions which correspond to bound states has no effect on the computed quantities, although it destroys the uniqueness of the solutions to (2.8).

THEOREM 5. *Let $\lambda_1 = \mu_1 + i\nu_1$, where μ_1 and ν_1 are real. Then*

$$\nu_1 = \frac{1}{4k}\left[1 + \frac{|(\varphi_1, \bar{\varphi}_0)|^2}{|(\varphi_1, \varphi_0)|^2}\right]. \tag{4.10}$$

PROOF.

$$2i\nu_1 = \frac{(\varphi_0, \varphi_1) - (\varphi_1, \varphi_0)}{|(\varphi_1, \varphi_0)|^2} = \frac{(A\varphi_1, \varphi_1) - (\varphi_1, A\varphi_1)}{|(\varphi_1, \varphi_0)|^2} = -\left\{\frac{(K\varphi_1, \varphi_1) - (\varphi_1, K\varphi_1)}{|(\varphi_1, \varphi_0)|^2}\right\}.$$

Using this and (2.5) we obtain (4.10). This theorem is the one-dimensional counterpart of the total cross-section theorem [7], [9]. It shows that the imaginary part of λ_1 is determined by the ratio of the reflected energy to the transmitted energy. If one assumes conservation of particles, theorem 5 can be proved analogously to the proof given by Schiff in the three-dimensional case [L. I. Schiff, Proc. Theor. Physics **11**, 288 (1954)]. The authors are indebted to Professor Schiff for this observation.

§ 5. The quadratic formulation

The variational principle stated in § 4 is in terms of a bilinear form and a bilinear constraint. We shall now show that the principle can be formulated as a quadratic problem. This will facilitate a heuristic understanding of the principle, and make clear the geometry of the "saddle" nature of the stationary point.

We introduce the space

$$\underline{H} = H \oplus H. \tag{5.1}$$

The elements of $\underline{H}$ will be of the form

$$\underline{\varphi} = [\varphi, \psi], \tag{5.2}$$

where φ and ψ are in H. If $\underline{\varphi}' = [\varphi', \psi']$ and $\underline{\varphi}'' = [\varphi'', \psi'']$ are in $\underline{H}$, their inner product will be given by

$$(\underline{\varphi}', \underline{\varphi}'') = (\varphi', \varphi'') + (\psi', \psi''). \tag{5.3}$$

We now state two lemmas. Their proofs are matters of routine computation, and hence will be omitted.

LEMMA 5. *Let $\underline{C}$ be a mapping of $\underline{H}$ into $\underline{H}$ defined by*

$$\underline{C}\,\underline{\varphi} = \underline{C}\,[\varphi, \psi] = [\overline{\psi}, \overline{\varphi}]. \tag{5.4}$$

Then $\underline{C}^2 = \underline{I}$ (identity), $(\underline{C}\,\underline{\varphi}, \underline{C}\,\underline{\psi}) = (\underline{\psi}, \underline{\varphi})$, $\underline{C}(\lambda\,\underline{\varphi}) = \overline{\lambda}\,\underline{C}\,\underline{\varphi}$. If $\underline{C}\,\underline{\varphi} = \underline{\varphi}$ and $\underline{C}\,\underline{\psi} = \underline{\psi}$, then $(\varphi, \psi) = (\psi, \varphi)$ and this inner product is real. If

$$\underline{\varphi}_0 = [\varphi_0, \overline{\varphi}_0], \quad \underline{\psi}_0 = [\varphi_0, -\overline{\varphi}_0], \tag{5.5}$$

then

$$\underline{C}\,\underline{\varphi}_0 = \underline{\varphi}_0, \quad \underline{C}\,\underline{\psi}_0 = -\underline{\psi}_0. \tag{5.6}$$

LEMMA 6. *Let $\underline{L}$ be the mapping of $\underline{H}$ into $\underline{H}$ defined by*

$$\underline{L}\,\underline{\varphi} = \underline{L}\,[\varphi, \psi] = [A\varphi, A\psi]. \tag{5.8}$$

Then, $\underline{L}$ is a linear operator; $\underline{L}^*$ the conjugate of $\underline{L}$ is given by

$$\underline{L}\,\underline{\varphi} = [A^*\varphi, A^*\psi], \tag{5.9}$$

and

$$\underline{L}^*\underline{C} = \underline{C}\,\underline{L}. \tag{5.10}$$

Now we define

$$\underline{J} = 2(\underline{L}\,\underline{\varphi}, \underline{C}\,\underline{\varphi}). \tag{5.11}$$

We shall consider the problem of making $\underline{J}$ stationary subject to the constraint,

$$(\underline{\varphi}, \underline{\varphi}_0)^2 - (\underline{\varphi}, \underline{\psi}_0)^2 = 4. \tag{5.12}$$

As in § 4, we introduce the Lagrange multiplier, λ, and set

$$\underline{J}_\lambda = \underline{J} - \lambda\{(\underline{\varphi}, \underline{\varphi}_0)^2 - (\underline{\varphi}, \underline{\psi}_0)^2\}. \tag{5.13}$$

Then,

$$\left\{\begin{aligned} \delta\,\underline{J}_\lambda &= 2(\underline{L}\,\underline{\varphi}, \underline{C}\,\delta\,\underline{\varphi}) + 2(\underline{L}\,\delta\,\underline{\varphi}, \underline{C}\,\underline{\varphi}) - \\ &\quad - 2\lambda\{(\underline{\varphi}, \underline{\varphi}_0)(\delta\,\underline{\varphi}, \underline{\varphi}_0) - (\underline{\varphi}, \underline{\psi}_0)(\delta\,\underline{\varphi}, \underline{\psi}_0)\} \\ &= 2(\delta\,\underline{\varphi}, 2\underline{C}\,\underline{L}\,\varphi - \overline{\lambda}\{(\underline{\varphi}_0, \underline{\varphi})\,\underline{\varphi}_0 - (\underline{\psi}_0, \underline{\varphi})\,\underline{\psi}_0\}). \end{aligned}\right. \tag{5.14}$$

This last equality makes use of (5.10) and lemma 5. Making use of lemma 5, we find that J_λ is stationary if and only if

$$2\underline{L}\,\underline{\varphi} = \lambda\{(\underline{\varphi}, \underline{\varphi}_0)\,\underline{\varphi}_0 + (\underline{\varphi}, \underline{\psi}_0)\,\underline{\psi}_0\}. \tag{5.15}$$

But

$$(\underline{\varphi}, \underline{\varphi}_0) = (\varphi, \varphi_0) + (\psi, \bar{\varphi}_0); \quad (\underline{\varphi}, \underline{\psi}_0) = (\varphi, \varphi_0) - (\psi, \bar{\varphi}_0). \tag{5.16}$$

Therefore

$$(\underline{\varphi}, \underline{\varphi}_0)\,\underline{\varphi}_0 + (\underline{\varphi}, \underline{\psi}_0)\,\underline{\psi}_0 = [2(\varphi, \varphi_0)\,\varphi_0, 2(\psi, \bar{\varphi}_0)\,\bar{\varphi}_0]. \tag{5.17}$$

Therefore, $\underline{J}_\lambda$ is stationary at $\underline{\varphi} = [\varphi, \psi]$ if and only if

$$A\varphi = \lambda(\varphi, \varphi_0)\,\varphi_0, \quad A^*\bar{\psi} = \bar{\lambda}(\bar{\psi}, \varphi_0)\,\varphi_0. \tag{5.18}$$

This is a consequence of (5.17), (5.15), lemma 6, and (2.3). We have proved the following theorem:

THEOREM 6. *There is a unique value of λ, λ_1, such that $\underline{J}_{\lambda_1}$ can be made stationary. λ_1 is the stationary value of J subject to the constraint* (5.12). *λ_1 is the same as in theorem* 3. *$\underline{\varphi}_1 = [\lambda_1\varphi_1, \lambda_1\psi_1]$ is a stationary point, and all other stationary points are of the form $\underline{\varphi}(\alpha) = \left[\alpha\lambda_1\varphi_1, \frac{1}{\alpha}\lambda_1\psi_1\right]$, $\alpha \neq 0$.*

A heuristic interpretation of theorem 6 is that we are finding that member of a family of quadric surfaces which is tangent to a fixed hyperboloid. The tangency takes place along a one (complex) parameter curve, $\underline{\varphi}(\alpha)$, in the surface (5.12). If we re-paramatrize the curve $\underline{\varphi}(\alpha)$ by letting $\alpha = e^\beta$, we see that the curve $\underline{\varphi}(\alpha)$ is the orbit of $\underline{\varphi}_1$ under the analytic group of operators $e^{\underline{B}\beta}$, where $\underline{B}\,\underline{\varphi} = \underline{B}[\varphi, \psi] = [\varphi, -\psi]$ [3].

§ 6. The imaginary part of $\underline{J}$

If $\underline{\varphi}$ is a stationary point, and if $\underline{\varphi} + \delta\underline{\varphi}$ is any point which lies on the surface (5.12), we define

$$\begin{cases} \Delta\underline{J} = 2(\underline{L}(\underline{\varphi} + \delta\underline{\varphi}), \underline{C}(\underline{\varphi} + \delta\underline{\varphi})) - 2(\underline{L}\,\underline{\varphi}, \underline{C}\,\underline{\varphi}) \\ \quad = 4(\underline{L}\,\underline{\varphi}, \underline{C}\,\delta\underline{\varphi}) + 2(\underline{L}\,\delta\underline{\varphi}, \underline{C}\,\delta\underline{\varphi}). \end{cases} \tag{6.1}$$

Since both $\underline{\varphi}$ and $\underline{\varphi} + \delta\underline{\varphi}$ satisfy (5.12), we have

$$2((\underline{\varphi}, \underline{\varphi}_0)\,\underline{\varphi}_0 + (\underline{\varphi}, \underline{\psi}_0)\,\underline{\psi}_0, \underline{C}\,\delta\underline{\varphi}) + (\delta\underline{\varphi}, \underline{\varphi}_0)^2 - (\delta\underline{\varphi}, \underline{\psi}_0)^2 = 0. \tag{6.2}$$

But at a stationary point (5.15) holds. Combining this with (6.1) and (6.2) we secure

$$\Delta\underline{J} = 2(\underline{L}\,\delta\underline{\varphi}, \underline{C}\,\delta\underline{\varphi}) - \lambda_1\{(\delta\underline{\varphi}, \underline{\varphi}_0)^2 - (\delta\underline{\varphi}, \underline{\psi}_0)^2\}. \tag{6.3}$$

LEMMA 7. *If $\underline{\varphi}$ is in $\underline{H}$, we define*

$$\underline{u} = \frac{1}{2}(\underline{\varphi} + \underline{C}\,\underline{\varphi}), \quad \underline{v} = \frac{1}{2i}(\underline{\varphi} - \underline{C}\,\underline{\varphi}). \tag{6.4}$$

Then

$$\underline{\varphi} = \underline{u} + i\underline{v}, \quad \underline{C}\,\underline{u} = \underline{u}, \quad \underline{C}\,\underline{v} = \underline{v}.$$

Also

$$\frac{1}{2}(\underline{\varphi}_0 + \underline{C}\,\underline{\varphi}_0) = \underline{\varphi}_0, \qquad \frac{1}{2i}(\underline{\psi}_0 - \underline{C}\,\underline{\psi}_0) = -i\,\underline{\psi}_0,$$

$$\frac{1}{2i}(\underline{\varphi}_0 - \underline{C}\,\underline{\varphi}_0) = 0, \qquad \frac{1}{2}(\underline{\psi}_0 + \underline{C}\,\underline{\psi}_0) = 0.$$

We define

$$\underline{u}_0 = \underline{\varphi}_0; \quad \underline{v}_0 = -i\,\underline{\psi}_0. \tag{6.5}$$

Then

$$\underline{C}\,\underline{u}_0 = \underline{u}_0, \quad \underline{C}\,\underline{v}_0 = \underline{v}_0.$$

PROOF. Direct computation, using lemma 5.

LEMMA 8. *Let*

$$\underline{G} = \frac{1}{2}(\underline{L} + \underline{L}^*), \quad \underline{F} = \frac{1}{2i}(\underline{L} - \underline{L}^*). \tag{6.6}$$

Then, $\underline{L} = \underline{G} + i\underline{F}$, $\underline{G}$ and $\underline{F}$ are self-adjoint operators, and $\underline{F}$ is positive semi-definite. Both $\underline{G}$ and $\underline{F}$ commute with $\underline{C}$, and if $\underline{C}\,\underline{\varphi} = \underline{\varphi}$, $\underline{C}\,\underline{\psi} = \underline{\psi}$, then $(\underline{F}\,\underline{\varphi}, \underline{\psi}) = (\underline{\psi}, \underline{F}\,\underline{\varphi}) = (\underline{\varphi}, \underline{F}\,\underline{\psi}) = (\underline{F}\,\underline{\psi}, \underline{\varphi})$, and this inner product is real. Similarly for $\underline{G}$.

PROOF. The first two assertions are obvious. To prove that $\underline{F}$ is positive semi-definite, we let $\underline{\varphi} = [\varphi, \psi]$. Then

$$\left\{\begin{aligned}
(\underline{F}\,\underline{\varphi}, \underline{\varphi}) &= \frac{1}{2i}(\underline{L}\,\underline{\varphi}, \underline{\varphi}) - \frac{1}{2i}(\underline{L}^*\underline{\varphi}, \underline{\varphi}) \\
&= \frac{1}{2i}([A\varphi, A\psi], [\varphi, \psi]) - \frac{1}{2i}([A^*\varphi, A^*\psi], [\varphi, \psi]) \\
&= \frac{1}{2i}\{((A - A^*)\varphi, \varphi) + ((A - A^*)\psi, \psi)\} \\
&= -\frac{1}{2i}\{((K - K^*)\varphi, \varphi) + ((K - K^*)\psi, \psi)\} \\
&= -\{(F\varphi, \varphi) + (F\psi, \psi)\}.
\end{aligned}\right. \tag{6.7}$$

Therefore, from lemma 3, $(\underline{F}\,\underline{\varphi}, \underline{\varphi}) \geq 0$. Now

$$\underline{C}\,\underline{F} = -\frac{1}{2i}(\underline{C}\,\underline{L} - \underline{C}\,\underline{L}^*) = \frac{1}{2i}(\underline{L}\,\underline{C} - \underline{L}^*\underline{C}) = \underline{F}\,\underline{C},$$

from lemmas 5 and 6. Similarly $\underline{C}\,\underline{G} = \underline{G}\,\underline{C}$. Finally,

$$(\underline{F}\,\underline{\varphi}, \underline{\psi}) = (\underline{F}\,\underline{C}\,\underline{\varphi}, \underline{C}\,\underline{\psi}) = (\underline{C}\,\underline{F}\,\underline{\varphi}, \underline{C}\,\underline{\psi}) = (\underline{\psi}, \underline{F}\,\underline{\varphi}),$$

from lemma 5. The other equalities follow from the self-adjointness of $\underline{F}$. Similarly for $\underline{G}$. This completes the proof.

Returning to ΔJ, setting

$$\delta\underline{\varphi} = \delta\underline{u} + i\,\delta\underline{v}, \tag{6.8}$$

where $\underline{C}\,\delta\underline{u}=\delta\underline{u}$, $\underline{C}\,\delta\underline{v}=\delta\underline{v}$ (lemma 7), we have

$$\underline{L}\,\delta\underline{\varphi}=(\underline{G}+i\underline{F})\,(\delta\underline{u}+i\,\delta\underline{v})=(\underline{G}\,\delta\underline{u}-\underline{F}\,\delta\underline{v})+i(\underline{F}\,\delta\underline{u}+\underline{G}\,\delta\underline{v}), \tag{6.9}$$

$$\left\{\begin{aligned}(\underline{L}\,\delta\underline{\varphi},\underline{C}\,\delta\underline{\varphi})&=\{\{(\underline{G}\,\delta\underline{u}-\underline{F}\,\delta\underline{v})+i\,(\underline{F}\,\delta\underline{u}+\underline{G}\,\delta\underline{v})\},\delta\underline{u}-i\,\delta\underline{v}\}\\&=\{(\underline{G}\,\delta\underline{u}-\underline{F}\,\delta\underline{v};\delta\underline{u})-(\underline{F}\,\delta\underline{u}+\underline{G}\,\delta\underline{v},\delta\underline{v})\}+\\&\quad+i\{(\underline{F}\,\delta\underline{u}+\underline{G}\,\delta\underline{v},\delta\underline{u})+(\underline{G}\,\delta\underline{u}-\underline{F}\,\delta\underline{v},\delta\underline{v})\}.\end{aligned}\right. \tag{6.10}$$

Now

$$(\delta\underline{\varphi},\underline{\varphi}_0)=(\delta\underline{u}+i\,\delta\underline{v},\underline{u}_0)=(\delta\underline{u},\underline{u}_0)+i(\delta\underline{v},\underline{u}_0), \tag{6.11}$$

$$(\delta\underline{\varphi},\underline{\psi}_0)=(\delta\underline{u}+i\,\delta\underline{v},i\,\underline{v}_0)=(\delta\underline{v},\underline{v}_0)-i(\delta\underline{u},\underline{v}_0). \tag{6.12}$$

Therefore,

$$\left\{\begin{aligned}(\delta\underline{\varphi},\underline{\varphi}_0)^2-(\delta\underline{\varphi},\underline{\psi}_0)^2&=\{(\delta\underline{u},\underline{u}_0)^2+(\delta\underline{u},\underline{v}_0)^2-(\delta\underline{v},\underline{u}_0)^2-(\delta\underline{v},\underline{v}_0)^2\}\\&\quad+2i\{(\delta\underline{u},\underline{u}_0)\,(\delta\underline{v},\underline{u}_0)+(\delta\underline{u},\underline{v}_0)\,(\delta\underline{v},\underline{v}_0)\}.\end{aligned}\right. \tag{6.13}$$

Then,

$$\left\{\begin{aligned}&\lambda_1\{(\delta\underline{\varphi},\underline{\varphi}_0)^2-(\delta\underline{\varphi},\underline{\psi}_0)^2\}\\&\quad=[\mu_1\{(\delta\underline{u},\underline{u}_0)^2+(\delta\underline{u},\underline{v}_0)^2-(\delta\underline{v},\underline{u}_0)^2-(\delta\underline{v},\underline{v}_0)^2\}-\\&\qquad-2\nu_1\{(\delta\underline{u},\underline{u}_0)\,(\delta\underline{v},\underline{u}_0)+(\delta\underline{u},\underline{v}_0)\,(\delta\underline{v},\underline{v}_0)\}]+\\&\qquad+i\,[\nu_1\{(\delta\underline{u},\underline{u}_0)^2+(\delta\underline{u},\underline{v}_0)^2-(\delta\underline{v},\underline{u}_0)^2-(\delta\underline{v},\underline{v}_0)^2\}+\\&\qquad+2\mu_1\{(\delta\underline{u},\underline{u}_0)\,(\delta\underline{v},\underline{u}_0)+(\delta\underline{u},\underline{v}_0)\,(\delta\underline{v},\underline{v}_0)\}].\end{aligned}\right. \tag{6.14}$$

LEMMA 9. *If* $\underline{J}=\underline{P}+i\,\underline{Q}$, $\underline{P}$ *and* $\underline{Q}$ *real, then* $\Delta\underline{J}=\Delta\underline{P}+i\Delta\underline{Q}$, *and*

$$\left\{\begin{aligned}\Delta\underline{P}&=2\{(\underline{G}\,\delta\underline{u},\delta\underline{u})-(\underline{G}\,\delta\underline{v},\delta\underline{v})-2(\underline{F}\,\delta\underline{u},\delta\underline{v})\}-\\&\quad-\mu_1\{(\delta\underline{u},\underline{u}_0)^2+(\delta\underline{u},\underline{v}_0)^2-(\delta\underline{v},\underline{u}_0)^2-(\delta\underline{v},\underline{v}_0)^2\}+\\&\quad+2\nu_1\{(\delta\underline{u},\underline{u}_0)\,(\delta\underline{v},\underline{u}_0)+(\delta\underline{u},\underline{v}_0)\,(\delta\underline{v},\underline{v}_0)\}.\end{aligned}\right. \tag{6.15}$$

$$\left\{\begin{aligned}\Delta\underline{Q}&=2\{(\underline{F}\,\delta\underline{u},\delta\underline{u})-(\underline{F}\,\delta\underline{v},\delta\underline{v})+2(\underline{G}\,\delta\underline{u},\delta\underline{v})\}-\\&\quad-\nu_1\{(\delta\underline{u},\underline{u}_0)^2+(\delta\underline{u},\underline{v}_0)^2-(\delta\underline{v},\underline{u}_0)^2-(\delta\underline{v},\underline{v}_0)^2\}-\\&\quad-2\mu_1\{(\delta\underline{u},\underline{u}_0)\,(\delta\underline{v},\underline{u}_0)+(\delta\underline{u},\underline{v}_0)\,(\delta\underline{v},\underline{v}_0)\}.\end{aligned}\right. \tag{6.16}$$

PROOF. (6.15) and (6.17) are obtained by combining (6.3), (6.10), and (6.14) and observing that all the inner products are real (lemma 5). The effect of lemma 9 is to express the increments of the real and imaginary parts of $\underline{J}$, as we move along the constraining surface, (5.12), away from a stationary point as real quadratic forms in $\delta\underline{u}$ and $\delta\underline{v}$. Since all the inner products which appear are real, we can think of the pair $\delta\underline{u}$, $\delta\underline{v}$ as being a point in the direct sum of two *real* Hilbert spaces. It should be observed however that these real Hilbert spaces are not spaces of real functions, but rather complicated spaces built up by using direct sums of complex function spaces and conjugation.

LEMMA 10. *Let $\underline{u}$ and $\underline{v}$ be obtained from the stationary point $\underline{\varphi}(\alpha)$ (theorem 6) according to* (6.4). Then

$$(6.17)\qquad (\underline{u},\underline{u}_0)=\frac{\alpha+\bar{\alpha}}{2}\left[\frac{|\alpha|^2+1}{|\alpha|^2}\right];\quad (\underline{u},\underline{v}_0)=-\frac{(\alpha-\bar{\alpha})}{2i}\left[\frac{|\alpha|^2+1}{|\alpha|^2}\right],$$

$$(6.18)\qquad (\underline{v},\underline{u}_0)=\frac{\alpha-\bar{\alpha}}{2i}\left[\frac{|\alpha|^2-1}{|\alpha|^2}\right];\quad (\underline{v},\underline{v}_0)=\frac{\alpha+\bar{\alpha}}{2}\left[\frac{|\alpha|^2-1}{|\alpha|^2}\right].$$

In general, if $\underline{\varphi}$ lies on (5.12), and $\underline{u}$ and $\underline{v}$ are given by (6.4), then

$$(6.19)\qquad (\underline{u},\underline{u}_0)^2+(\underline{u},\underline{v}_0)^2-(\underline{v},\underline{u}_0)^2-(\underline{v},\underline{v}_0)^2=4,$$

$$(6.20)\qquad (\underline{u},\underline{u}_0)(\underline{v},\underline{u}_0)+(\underline{u},\underline{v}_0)(\underline{v},\underline{v}_0)=0.$$

PROOF. We shall carry out the computation for the first equation in (6.17).

$$\begin{aligned}(\underline{u},\underline{u}_0)&=\tfrac{1}{2}(\underline{\varphi}(\alpha)+\underline{C}\,\underline{\varphi}(\alpha),\underline{\varphi}_0)=\tfrac{1}{2}(\underline{\varphi}(\alpha),\underline{\varphi}_0)+\tfrac{1}{2}(\underline{\varphi}(\alpha))\\&=\tfrac{1}{2}\left(\left[\alpha\lambda_1\varphi_1,\tfrac{1}{\alpha}\lambda_1\psi_1\right],[\varphi_0,\bar{\varphi}_0]\right)+\tfrac{1}{2}\left([\varphi_0,\bar{\varphi}_0],\left[\alpha\lambda_1\varphi_1,\tfrac{1}{\alpha}\lambda_1\psi_1\right]\right)\\&=\tfrac{1}{2}\left\{\alpha+\tfrac{1}{\alpha}+\bar{\alpha}+\tfrac{1}{\alpha}\right\}=\frac{\alpha+\bar{\alpha}}{2}\left[\frac{|\alpha|^2+1}{|\alpha|^2}\right].\end{aligned}$$

The remaining equations in (6.17) and (6.18) are derived in a similar manner. As for (6.19) and (6.20), they are obtained analogously to (6.13).

THEOREM 7. *Let $\underline{u}$ and $\underline{v}$ be obtained by* (6.4) *from the stationary point $\underline{\varphi}(\alpha)$. Let $\delta\underline{u}$ and $\delta\underline{v}$ be similarly obtained from $\delta\underline{\varphi}$, where $\underline{\varphi}(\alpha)+\delta\underline{\varphi}$ satisfies* (5.12). *Then*

(a) *If $|\alpha|\neq 1$ there exists an $\varepsilon<0$, such that in the weak neighborhood of $\underline{\varphi}(\alpha)$ defined by $(\delta\underline{u},\underline{u}_0)^2+(\delta\underline{u},\underline{v}_0)^2<\varepsilon$, $(\delta\underline{v},\underline{u}_0)^2+(\delta\underline{v},\underline{v}_0)^2<\varepsilon$, the increment in the imaginary part of $\underline{J}$ satisfies*

$$(6.21)\qquad \Delta Q]_{\delta\underline{u}=0}=-(\underline{F}\,\delta\underline{v},\delta\underline{v})\leqq 0,$$

$$(6.22)\qquad \Delta Q]_{\delta\underline{v}=0}=(\underline{F}\,\delta\underline{u},\delta\underline{u})\geqq 0.$$

(b) *If $|\alpha|=1$, then $\Delta Q]_{\delta\underline{u}=0}=-(\underline{F}\,\delta\underline{v},\delta\underline{v})\leqq 0$.*

PROOF. Expanding (6.2) with the help of (6.13), we obtain

$$(6.23)\qquad \begin{aligned}&2\{(\underline{u},\underline{u}_0)(\delta\underline{u},\underline{u}_0)+(\underline{u},\underline{v}_0)(\delta\underline{u},\underline{v}_0)-(\underline{v},\underline{u}_0)\delta\underline{v},\underline{u}_0)-(\underline{v},\underline{v}_0)(\delta\underline{v},\underline{v}_0)\}+\\&+(\delta\underline{u},\underline{u}_0)^2+(\delta\underline{u},\underline{v}_0)^2-(\delta\underline{v},\underline{u}_0)^2-(\delta\underline{v},\underline{v}_0)^2=0,\end{aligned}$$

$$(6.24)\qquad \begin{aligned}&(\underline{u},\underline{u}_0)(\delta\underline{v},\underline{u}_0)+(\underline{u},\underline{v}_0)(\delta\underline{v},\underline{v}_0)+(\underline{v},\underline{u}_0)(\delta\underline{u},\underline{u}_0)+(\underline{v},\underline{v}_0)(\delta\underline{u},\underline{v}_0)+\\&+(\delta\underline{u},\underline{u}_0)(\delta\underline{v},\underline{u}_0)+(\delta\underline{u},\underline{v}_0)(\delta\underline{v},\underline{v}_0)=0.\end{aligned}$$

(a) Suppose $\delta\underline{u}=0$. Then $\delta\underline{v}$ must satisfy

$$(6.25)\qquad 2\{(\underline{v},\underline{u}_0)(\delta\underline{v},\underline{u}_0)+(\underline{v},\underline{v}_0)(\delta\underline{v},\underline{v}_0)\}+(\delta\underline{v},\underline{u}_0)^2+(\delta\underline{v},\underline{v}_0)^2=0$$

$$(6.26)\qquad (\underline{u},\underline{u}_0)(\delta\underline{v},\underline{u}_0)+(\underline{u},\underline{v}_0)(\delta\underline{v},\underline{v}_0)=0.$$

From lemma 10, it is clear that not both $(\underline{u}, \underline{u}_0) = 0$ and $(\underline{v}, \underline{v}_0) = 0$. Therefore, from the homogeneity of (6.20) and (6.26), we have

$$(\delta \underline{v}, \underline{u}_0)(\underline{v}, \underline{v}_0) - (\delta \underline{v}, \underline{v}_0)(\underline{v}, \underline{u}_0) = 0. \tag{6.27}$$

But (6.27) and (6.25) imply by Cramer's Rule that either

$$(\delta \underline{v}, \underline{u}_0) = (\delta \underline{v}, \underline{v}_0) = 0, \tag{6.28}$$

or

$$(\delta \underline{v}, \underline{u}_0) = -2(\underline{v}, \underline{u}_0); \quad (\delta \underline{v}, \underline{v}_0) = -2(\underline{v}, \underline{v}_0). \tag{6.29}$$

If (6.29) holds, then, from (6.18), we have

$$(\delta \underline{v}, \underline{u}_0)^2 + (\delta \underline{v}, \underline{v}_0)^2 = 4\left(|\alpha| - \frac{1}{|\alpha|}\right)^2. \tag{6.30}$$

Let $\delta \underline{v} = 0$. Then $\delta \underline{u}$ must satisfy

$$2\{(\underline{u}, \underline{u}_0)(\delta \underline{u}, \underline{u}_0) + (\underline{u}, \underline{v}_0)(\delta \underline{u}, \underline{v}_0)\} + (\delta \underline{u}, \underline{u}_0)^2 + (\delta \underline{u}, \underline{v}_0)^2 = 0, \tag{6.31}$$

$$(\underline{v}, \underline{u}_0)(\delta \underline{u}, \underline{u}_0) + (\underline{v}, \underline{v}_0)(\delta \underline{u}, \underline{v}_0) = 0. \tag{6.32}$$

Since it follows from (6.18) that not both $(\underline{v}, \underline{u}_0) = 0$ and $(\underline{v}, \underline{v}_0) = 0$, we find that here $\delta \underline{u}$ must satisfy either

$$(\delta \underline{u}, \underline{u}_0) = (\delta \underline{u}, \underline{v}_0) = 0 \tag{6.33}$$

or

$$(\delta \underline{u}, \underline{u}_0)^2 + (\delta \underline{u}, \underline{v}_0)^2 = 4\left(|\alpha| + \frac{1}{|\alpha|}\right)^2. \tag{6.34}$$

If we let $\varepsilon \leq 4\left(|\alpha| - \frac{1}{|\alpha|}\right)^2$, then in the neighborhood determined by this ε, when $\delta \underline{u} = 0$ (6.30) is excluded, and when $\delta \underline{v} = 0$, (6.34) is excluded. Therefore, from (6.16), (6.28), and (6.33) in this neighborhood (6.21) and (6.22) are valid.

(b) $|\alpha| = 1$. The discussion above applies through (6.30). But if $|\alpha| = 1$, both (6.28) and (6.30) imply $(\delta \underline{v}, \underline{v}_0) = (\delta \underline{v}, \underline{u}_0) = 0$, so the theorem is proved. We can now interpret theorem 7 in terms of the principle formulated in § 4.

COROLLARY. *Let* $\varphi_s = \alpha \lambda_1 \varphi_1$, $\psi_s = \frac{1}{\alpha} \lambda_1 \psi_1$ *be a pair of functions which make* $J = (A\varphi, \underline{\psi})$ *stationary subject to the constraint* $(\varphi, \varphi_0)(\varphi_0, \overline{\psi}) = 1$. *Then if* φ *and* ψ *are varied about* φ_s, ψ_s *subject to the constraint, then*

(a) *if* $|\alpha| \neq 1$, *and if* $\varphi + \overline{\psi}$ *is held constant, then* Im J *will have a local maximum at* φ_s, ψ_s. *If* $\varphi - \overline{\psi}$ *is held constant,* Im J *will have a local minimum at* φ_s, ψ_s.

(b) *if* $|\alpha| = 1$, *and if* $\varphi + \overline{\psi}$ *is held constant, then* Im J *will have a global maximum.*

§ 7. The variational principle for the reflection amplitude

The complex reflection amplitude in one dimensional quantum scattering is given by

$$\frac{1}{\mu_1} = \int_{-\infty}^{\infty} \omega(x) V(x) e^{+ikx} dx \tag{7.1}$$

where $\omega(x)$ is the solution of the integral equation (3.4). If, as before φ_0 is defined by (2.4), the appropriate integral equations corresponding to the set (2.8) are

$$A\varphi = \varphi_0, \quad A^*\bar{\psi} = \bar{\varphi}_0. \tag{7.2}$$

As noted in the proof of theorem 1, these equations have a unique solution in H. Moreover in view of lemma 1, here $\bar{\psi} = \bar{\varphi}$. The reciprocity relation corresponding to (2.9) reduces here to the trivial statement that

$$(\varphi_0, \bar{\varphi}) = (\varphi, \bar{\varphi}_0).$$

Now if φ_1 once again is the unique solution to the first of the set (7.2), then

$$\frac{1}{\mu_1} = (\varphi_1, \bar{\varphi}_0)$$

so that here we may treat (7.2) instead of (3.4).

The variational principle for the determination of $1/\mu_1$ consists in this case of making

$$J^R = (A\varphi, \bar{\varphi}), \tag{7.3}$$

stationary subject to the constraint

$$(\varphi, \bar{\varphi}_0) = 1. \tag{7.4}$$

To accomplish this, we introduce the Lagrange multiplier μ, and set

$$J^R_\mu = J^R - 2\mu(\varphi, \bar{\varphi}_0)$$

and obtain the result that J^R is stationary if and only if

$$A\varphi = \mu\varphi_0.$$

This last equation corresponds to the set (4.5). Here, as in the set (7.2), the second equation is merely the conjugate of this and consequently it may be omitted. Equation (7.4) is simpler in appearance than the set (4.5) since here we may use a linear rather than a quadratic constraint. Actually, if we think of the variational principles (5.11), (5.12), or (5.13) abstractly; that is, independent of the interpretation given to them in § 5, (7.3) and (7.4) can be thought of as special cases of (5.11) and (5.12) obtained by replacing φ_0 by $C\varphi_0$, setting ψ_0 equal to zero and interpreting C as usual conjugation. While a more unified treatment would result if (5.13) were similarly specialized, the simplifications resulting from the use of a linear rather than a quadratic restraint make the use of (7.4) preferable.

In order to obtain a theorem analogous to theorem 3, we now make the additional assumption that $(\varphi_1, \bar{\varphi}_0) \neq 0$. Physically this assumption is of course equivalent to assuming that the potential barrier is at least partially reflecting. It is moreover, not equivalent to assuming that the potential barrier is not transparent [i.e., that $(\varphi_1, \varphi_0) \neq 0$]. It is interesting to note

that the existence of partial reflection implies the existence of partial transmission. More explicitly we have the following:

LEMMA 11. *Let φ_1 be the unique solution to* (7.2). *Then* $(\varphi_1, \bar{\varphi}_0) \neq 0$ *implies that* $(\varphi_1, \varphi_0) \neq 0$.

PROOF. The argument immediately below equation (3.15) shows that $(\varphi_1, \varphi_0) = 0$ implies $(\varphi_1, \bar{\varphi}_0) = 0$. The lemma follows by logical equivalence.

THEOREM 8. *There exists exactly one value of μ, μ_1, such that J^R can be made stationary. μ_1 is equal to the value of J^R when J^R is made stationary subject to the constraint* (7.4). *Simultaneously,*

$$\mu_1 = \frac{1}{(\varphi_1, \bar{\varphi}_0)}.$$

Since the proof is nearly identical to that of theorem 3, we will content ourselves with remarking that if φ_s satisfies (7.5) it must be of the form $\varphi_s = \mu\varphi_1$, since φ_1 is the unique solution of (7.2). Moreover since φ_s must also satisfy (7.4), it follows that $\mu(\varphi_1, \bar{\varphi}_0) = 1$. Now (7.5) implies that

$$(A\,\varphi, \bar{\varphi}) = \mu(\varphi_0, \bar{\varphi}) = \mu.$$

Finally, in analogy to the corollary of theorem 7, we have

THEOREM 9. *Let $\varphi_s = \mu\varphi_1$ be any function which makes $J^R = (A\varphi, \bar{\varphi})$ stationary subject to the constraint $(\varphi, \bar{\varphi}_0) = 1$, then*

(a) *if $\varphi + \bar{\varphi}$ is held constant,* Im J^R *will have a global minimum at φ_s;*

(b) *if $\varphi - \bar{\varphi}$ is held constant,* Im J^R *will have a global maximum at φ_s;*

PROOF. Since here we are dealing with the linear constraint (7.4), the second variations of J^R and J^R_u are identical for any function φ_s which makes them stationary. The nature of the stationary point is therefore determined completely by the second variation $\delta^2 J^R$ which is

$$\delta^2 J^R = (A\,\delta\varphi, \delta\bar{\varphi}).$$

From this we find from (2.3),

$$\begin{aligned} 2i \operatorname{Im} \delta^2 J^R &= (A\,\delta\varphi, \delta\bar{\varphi}) - \overline{(A\,\delta\varphi, \delta\bar{\varphi})} \\ &= (A\,\delta\varphi, \delta\bar{\varphi}) - (A^*\,\delta\bar{\varphi}, \delta\varphi) \\ &= (\delta\varphi, \delta\bar{\varphi}) - (\delta\bar{\varphi}, \delta\varphi) + (K\,\delta\varphi, \delta\bar{\varphi}) - (K^*\,\delta\bar{\varphi}, \delta\varphi). \end{aligned}$$

Letting $\varphi = u + iv$ where u and v are real so that

$$\delta\varphi = \delta u + i\,\delta v$$

and using (2.6) we find that

$$2 \operatorname{Im} \delta^2 J^R = 2(\delta u, \delta v) + (F\,\delta u, \delta u) - (F\,\delta v, \delta v) + 2(G\,\delta u, \delta v).$$

The theorem then follows from lemma 3 since F is negative semi-definite.

In conclusion it should be remarked that we have been unable to find useful analog of theorem 5 for this case.

References

[1] DOLPH, C. L.: The Mathematician Grapples with Linear Problems Associated with the Radiation Condition. Trans. Inst. Radio Engineers PGAP, AP—4, (1956). — [2] DUNFORD, N.: Spectral operators. Pac. J. Math. **4**, 321—354 (1954). — [3] HILLE, E.: Functional Analysis and Semi-Groups. Amer. Math. Soc., Colloquium Publ. **31** (1948). — [4] JONES, D. S.: A Critique of the Variational Method. Trans. Inst. Radio Engineers PGAP, AP—4, (1956). — [5] KATO, T.: Note on Schwinger Variational Principle. Progress of Theoretical Physics (Japan) **6**, 295—305 (1951). — [6] MARCUVITZ, N.: Recent Developments in Wave Propagation, Notes, Part III-D. New York University 1953. — [7] MORSE and FESHBACH: Methods of Theoretical Physics, Vol II, Chapters 9, 11. New York: McGraw Hill 1953. — [8] RIESZ, F., and B. Sz. NAGY: Lecons d'Analyse Fonctionnelle. Budapest 1952. — [9] SCHWINGER, J., and H. LEVINE: On the Theory of Diffraction by an Aperture in an Infinite Plane Screen. I. Physical Review **74**, 958—974 (1948).

Engineering Res. Inst., University of Michigan, Ann Arbor, Mich. (U.S.A.)

(Eingegangen am 9. Januar 1956)

AVAKUMOVIĆ, V. G.
Math. Zeitschr. Bd. 65, S. 327—344 (1956)

Über die Eigenfunktionen auf geschlossenen Riemannschen Mannigfaltigkeiten

Von
VOJISLAV G. AVAKUMOVIĆ

1. Einleitung. *1.1.* Sei D ein beschränktes Gebiet des N-dimensionalen Euklidischen Raumes. Der Rand $\overline{D}$ von D sei stückweise glatt. P bezeichne Punkte aus D und λ_n bzw. $\Phi_n(P)$ die Eigenwerte bzw. die auf D orthonormierten Eigenfunktionen von $(\Delta+\lambda)u=0$ innerhalb D und $u=0$ längs $\overline{D}$.

Bezüglich λ_n und $\Phi_n(P)$ gelten folgende zwei fundamentale asymptotische Formeln:

$$\sum_{\lambda_n \leq x} 1 = F C_N x^{N/2} + o(x^{N/2}) \tag{1.1}$$

und

$$\sum_{\lambda_n \leq x} \Phi_n^2(P) = C_N x^{N/2} + o(x^{N/2}). \tag{1.2}$$

Dabei ist $C_N = 1/(2\sqrt{\pi})^N \, \Gamma(N/2+1)$ und F bezeichnet den Rauminhalt von D. Die erste dieser Formeln wurde von Herrn WEYL [14] im Jahre 1911 und die zweite von Herrn CARLEMAN [4] im Jahre 1934 bewiesen.

Herr COURANT [5] hat in (1.1) das $o\left(x^{\frac{N}{2}}\right)$ bis auf $O\left(x^{\frac{N-1}{2}} \lg x\right)$ herabgedrückt; der Verf. [2] und unabhängig davon Herr LEVITAN [9a] haben gezeigt, daß in (1.2) das $o\left(x^{\frac{N}{2}}\right)$ bis auf $O\left(x^{\frac{N-1}{2}}\right)$ verschärft werden kann (welches Ergebnis bestmöglich ist).

Schließlich haben die Herren MINAKSHISUNDARAM und PLEIJEL [10] die Formeln (1.1) und (1.2) auf RIEMANNsche Mannigfaltigkeiten übertragen.

1.2. Sei V eine geschlossene, zusammenhängende, kompakte und beliebig oft differenzierbare RIEMANNsche Mannigfaltigkeit[1]). $P, Q, \Pi, \ldots, X, Y$ bezeichnen Punkte aus V, x^i $(i=1, 2, \ldots, N)$ die lokalen Koordinaten und $g_{ik}(x^1, x^2, \ldots, x^N)$ die Komponenten des metrischen Tensors. Ferner bezeichne Δ den BELTRAMIschen Operator

$$\Delta = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^i}\left(\sqrt{g}\, g^{ik} \frac{\partial}{\partial x^k}\right),$$

wobei $g^{ik} g_{kj} = \delta_j^i$ und $g = \det(g_{ij})$ ist.

[1]) Aus dem Folgenden (insbesondere dem Beweis des Hilfssatzes 1) ist ersichtlich, daß die Voraussetzung „V ist kompakt und beliebig oft differenzierbar" durch die allgemeinere „$V \in C^6$" ersetzt werden kann. Dabei ist, wie üblich, „$V \in C^r$" folgendermaßen erklärt: 1. V ist eine kompakte Mannigfaltigkeit. 2. Sind x^i und x'^i zwei Lokalkoordinatensysteme irgendeines Punktes aus V, so haben $x^i(x'^j)$ und $x'^i(x^j)$ stetige partielle Ableitungen r-ter Ordnung. 3. Die Komponenten g^{ik} der eingeprägten Metrik haben in bezug auf Lokalkoordinaten stetige partielle Ableitungen der Ordnung $r-1$.

Da V geschlossen ist so hat, wie bekannt, die Aufgabe $(\Delta + \lambda)u = 0$ dann und nur dann auf V reguläre Lösungen, wenn λ gleich einem der Eigenwerte

$$\lambda_0, \lambda_1, \lambda_2, \ldots$$

ist. Dabei ist $\lambda_0 = 0$ und $\lambda_n \leq \lambda_{n+1} \to \infty$; jedem λ_n entspricht eine und nur eine orthonormierte Eigenfunktion. Diese seien mit

$$\Phi_0(P), \Phi_1(P), \Phi_2(P), \ldots$$

bezeichnet. Dabei ist $\Phi_0(P) = 1/\sqrt{F}$ mit $F =$ Rauminhalt von V.

Nach MINAKSHISUNDARAM und PLEIJEL gelten, wie schon erwähnt, folgende Formeln:

$$\sum_{\lambda_n \leq x} 1 = F C_N x^{N/2} + o(x^{N/2})$$

und, falls

$$\Theta(P; x) = \sum_{\lambda_n \leq x} \Phi_n^2(P)$$

gesetzt wird,

$$\Theta(P; x) = C_N x^{\frac{N}{2}} + o\left(x^{\frac{N}{2}}\right).$$

In der vorliegenden Arbeit werde ich zeigen, daß die in 1.1 erwähnten Verschärfungen auch auf V gelten und daß die COURANTsche Abschätzung sogar noch verschärft werden kann. Um dabei das Wesentliche hervortreten zu lassen und technische Schwierigkeiten zu vermeiden, beschränke ich mich auf den Fall $N = 3$[2]).

1.3. Satz 1. *Es mögen V und $\Theta(P; x)$ dieselbe Bedeutung wie in 1.2 haben und es sei $N = 3$. Dann gilt*

(I) $$\Theta(P; x) = \frac{1}{6\pi^2} x^3 + O(x),$$

und zwar gleichmäßig in bezug auf P.

Wird $\Theta(P; x)$ über V integriert, so folgt aus (I) der

Satz 2. *Unter den Voraussetzungen des Satzes 1 gilt*

(II) $$\sum_{\lambda_n \leq x} 1 = \frac{F}{6\pi^2} x^3 + O(x).$$

Dabei hat F dieselbe Bedeutung wie in 1.2.

Der hier gegebene Beweis von (I) mit einem von P unabhängigen $O(x)$-Glied beruht wesentlich auf der Voraussetzung, daß V unberandet ist.

Schließlich werde ich am Beispiel der dreidimensionalen Kugelfläche zeigen, daß die gewonnenen Ergebnisse nicht verbessert werden können. Es gilt also der

Satz 3. *Die $O(x)$-Abschätzungen in* (I) *und* (II) *sind bestmöglich.*

---

[2]) Im Falle $N \neq 3$ ist es zweckmäßig anstatt der Schwingungsgleichung die Wärmeleitungsgleichung $\left(\Delta - \frac{\partial}{\partial t}\right) u = 0$ zu verwenden und folglich mit $\sum \Phi_n(P)\, \Phi_n(Q)\, e^{-\lambda_n t}$ anstatt mit $\sum \Phi_n(P)\, \Phi_n(Q)/\lambda_n^{2[N/4]+1}\, (\lambda_n + \lambda)$ zu arbeiten.

1.4. Im folgenden bezeichnen $m_1, m_2, m_3, \ldots$ feste, nur von V abhängige Zahlen.

Die Zahl R ist folgendermaßen erklärt: Da V geschlossen und kompakt ist, gibt es eine von P unabhängige feste Zahl $R>0$, derart, daß die geodätische Kugel um P vom Durchmesser R ganz im Existenzbereich jedes Normalkoordinatensystems mit Zentrum P liegt.

Den Satz 1 werde ich aus folgendem Satz herleiten.

Satz A. *Es gibt zwei Funktionen $R(P;x)$ und $w(P)$ mit*

$$|R(P;x)| \leq m_1 \quad \text{und} \quad |w(P)| \leq m_2 \tag{1.3}$$

und eine von P unabhängige feste Zahl ϱ_0 mit $0<2\varrho_0 \leq R$, so daß folgendes gilt: Wird

$$A(P;x) = \sum_{\lambda_1 \leq \lambda_n \leq x} \frac{\Phi_n^2(P)}{\lambda_n} - \frac{\sqrt{x}}{2\pi^2} - w(P) + R(P;x)$$

gesetzt, so konvergiert

$$\varphi(P;s) = \int_0^\infty e^{-sv} A(P;v^2)\, dv$$

für $\Re(s)>0$[3]) *und ist daselbst einschließlich $|s|<\frac{\varrho_0}{2}$ eine reguläre analytische Funktion und genügt für $|s| \leq \frac{\varrho_0}{3}$ der Ungleichung*

$$|\varphi(P;s)| \leq m_3. \tag{1.4}$$

1.5. Wegen (1.3) und

$$A(P;y^2) - A(P;x^2) = \sum_{x^2 < \lambda_n \leq y^2} \frac{\Phi_n^2(P)}{\lambda_n} - \frac{1}{2\pi^2}(y-x) + \left(R(P;y^2) - R(P;x^2)\right)$$

ist

$$A(P;y^2) - A(P;x^2) \geq -m_5 \quad \text{für } x \leq y \leq x+1. \tag{1.5}$$

Aus dem regulären Verhalten von $\varphi(P;s)$ im Punkte $s=0$ und der Ungleichung (1.5) ergibt sich auf bekannte Weise (Verallgemeinerung des FATOU-RIESZschen O-Satzes; vgl. [8] und [9] sowie den Beweis in *1.6*)

$$A(P;x) = O(1) \quad (x \to \infty). \tag{1.6}$$

Wegen (1.4) und (1.5) gilt (1.6) gleichmäßig in bezug auf P, so daß wegen (1.3)

$$\left| \sum_{\lambda_1 \leq \lambda_n \leq x} \frac{\Phi_n^2(P)}{\lambda_n} - \frac{\sqrt{x}}{2\pi^2} \right| \leq m_6 \tag{1.7}$$

ist. Setzen wir

$$\Theta_1(P;x) = \sum_{\lambda_1 \leq \lambda_n \leq x} \frac{\Phi_n^2(P)}{\lambda_n},$$

so ist

$$\Theta(P;x) = \int_0^x t\, d\{\Theta_1(P;t)\} = x\,\Theta_1(P;x) - \int_0^x \Theta_1(P;t)\, dt,$$

woraus sich auf Grund von (1.7) die Beziehung (I) ergibt.

[3]) $\Re(s)$ bezeichnet den reellen Teil von S.

1.6. Der Vollständigkeit wegen wollen wir (1.6) kurz beweisen. Dabei stütze ich mich auf die in 4.1 zu beweisende Formel $A(P;x)=O\left(\sqrt{x}\right)$, $x\to\infty$.

a) Wir setzen (vgl. [8] und [9])

$$p(x)=\begin{cases}\frac{8}{3}-4x^2+4x^3 & \text{für } 0\leq x\leq\frac{1}{2},\\ \frac{4}{3}(1-x)^3 & \text{für } \frac{1}{2}\leq x\leq 1,\\ 0 & \text{für } x\geq 1,\end{cases}$$

so daß

$$\int_{-\frac{\varrho_0}{3}}^{+\frac{\varrho_0}{3}} e^{iyt}\,p\left(\frac{3|t|}{\varrho_0}\right)dt=\frac{\varrho_0}{6}\left(\frac{\sin\varrho_0 y/12}{\varrho_0 y/12}\right)^4$$

ist. Mit $\varepsilon>0$ bilden wir

$$H_\varepsilon(P;y)=\int_{-\frac{\varrho_0}{3}}^{+\frac{\varrho_0}{3}} e^{iyt}\,p\left(\frac{3|t|}{\varrho_0}\right)\varphi(P;\varepsilon+it)\,dt.$$

Da $\varphi(P;\varepsilon+it)$ für $\varepsilon>0$ als LAPLACE-Integral darstellbar ist, so ist (Satz von FUBINI)

$$(1.8)\qquad H_\varepsilon(P;y)=\int_0^\infty e^{-\varepsilon v}A(P;v^2)\,dv\int_{-\frac{\varrho_0}{3}}^{+\frac{\varrho_0}{3}} e^{i(y-v)t}\,p\left(\frac{3|t|}{\varrho_0}\right)dt$$

$$=\frac{\varrho_0}{6}\int_0^\infty e^{-\varepsilon v}A(P;v^2)\left(\frac{\sin\varrho_0(y-v)/12}{\varrho_0(y-v)/12}\right)^4 dv.$$

b) Da $\varphi(P;s)$ für $|s|\leq\frac{\varrho_0}{3}$ regulär ist, existiert

$$H(P;y)=\lim_{\varepsilon\to 0}H_\varepsilon(P;y)$$

und ist

$$=\int_{-\frac{\varrho_0}{3}}^{+\frac{\varrho_0}{3}} e^{iyt}\,p\left(\frac{3|t|}{\varrho_0}\right)\varphi(P;it)\,dt.$$

Wegen (1.4) und $|p(x)|\leq\frac{8}{3}$ ist also

$$(1.9)\qquad |H(P;y)|\leq\tfrac{4}{3}\varrho_0 m_3.$$

c) Wegen $A(P;x)=O\left(\sqrt{x}\right)$ kann man im rechtsstehenden Integral (1.8) den Grenzübergang $\varepsilon\to 0$ vor und hinter dem Integralzeichen vertauschen, so daß

$$(1.10)\qquad H(P;y)=\frac{\varrho_0}{6}\int_0^\infty A(P;v^2)\left(\frac{\sin\varrho_0(y-v)/12}{\varrho_0(y-v)/12}\right)^4 dv$$

wird.

d) Aus (1.5), (1.9) und (1.10) folgt auf bekannte und elementare Weise $A(P; x) = O(1)$ für $x \to \infty$ (vgl. G. H. HARDY, Divergent Series, S. 305), und zwar [wegen (1.5) und (1.9)] gleichmäßig in bezug auf P.

2. Die Parametrix. *2.1.* Nach MINAKSHISUNDARAM und PLEIJEL [*10*] kann man eine Parametrix von

$$(\Delta - \lambda)\, u = 0 \tag{2.1}$$

wie folgt konstruieren: Der metrische Tensor sei o.B.d.A. auf Normalkoordinaten bezogen und $r = r_{PQ}$ bezeichne die geodätische Entfernung der Punkte P und Q. Dann gilt für jede nur von r abhängige Funktion $\Phi(r)$ und jede Funktion u die Formel

$$\Delta \Phi u = u\left(\frac{d^2\Phi}{dr^2} + \frac{2}{r}\frac{d\Phi}{dr} + \frac{d(\lg\sqrt{g})}{dr}\frac{d\Phi}{dr}\right) + 2\frac{du}{dr}\frac{d\Phi}{dr} + \Phi\,\Delta u. \tag{2.2}$$

Wird $u_\nu = u_\nu(P, Q)$ $(\nu = 0, 1, 2, \ldots)$ mittels

$$r\frac{du_\nu}{dr} + \frac{r}{2}\frac{d(\lg\sqrt{g})}{dr}u_\nu + \nu u_\nu = \Delta u_{\nu-1};\ u_0(P, P) = 1,\ u_{-1} \equiv 0 \tag{2.3}$$

definiert und

$$K_m(P, Q; -\lambda) = \sum_{\nu=0}^{m} u_\nu(P, Q)\, k_\nu(P, Q; -\lambda)$$

mit

$$k_\nu(P, Q; -\lambda) = \frac{1}{(2\sqrt{\pi})^3}\int_0^\infty e^{-\lambda t - \frac{r^2}{4t}} t^{\nu - \frac{3}{2}}\, dt$$

gesetzt, so ist wegen (2.2)

$$(\Delta - \lambda)\, K_m(P, Q; -\lambda) = k_m(P, Q; -\lambda)\, \Delta u_m(P, Q). \tag{2.4}$$

Da $K_m(P, Q; -\lambda)$ für $m \geq 0$ dieselbe Singularität wie die GREENsche Funktion von (2.1) besitzt, ist $K_m(P, Q; -\lambda)$ eine lokale Parametrix von (2.1).

Für jedes $\varrho > 0$ gibt es ein viermal stetig differenzierbares $\eta = \eta(r)$, so daß

$$\eta(r) = \begin{cases} 1 & \text{für } 0 \leq r \leq \varrho \\ 0 & \text{für } r \geq 2\varrho \end{cases}$$

und

$$|\eta(r)| \leq 1,\ \left|\frac{d\eta(r)}{dr}\right| \leq \frac{\text{const}}{r} \quad \text{und} \quad \left|\frac{d^2\eta(r)}{dr^2}\right| \leq \frac{\text{const}}{r^2}. \tag{2.5}$$

Für jedes feste $0 < 2\varrho \leq R$ und $m \geq 0$ ist

$$\Gamma_m(P, Q; -\lambda) = \eta(r_{PQ})\, K_m(P, Q; -\lambda)$$

eine auf $V \times V$ erklärte Parametrix im großen.

2.2. Auf Grund von (2.2) und (2.4) ist

$$(2.6)\quad \begin{cases} \Omega_m(P,Q;-\lambda) = (\Delta-\lambda)\,\Gamma_m(P,Q;-\lambda) \\ \qquad = k_m(P,Q;-\lambda)\,\Delta u_m(P,Q) & \text{für } 0\leq r\leq \varrho, \\ \qquad = \left\{\frac{d^2\eta}{dr^2}+\frac{2}{r}\frac{d\eta}{dr}+\frac{d(\lg\sqrt{g})}{dr}\frac{d\eta}{dr}\right\}K_m(P,Q;-\lambda)+ \\ +2\frac{d\eta}{dr}\frac{dK_m(P,Q;-\lambda)}{dr}+\eta\,k_m(P,Q;-\lambda)\,\Delta u_m(PQ) & \text{für } \varrho\leq r\leq 2\varrho \end{cases}$$

und

$$= 0 \qquad \text{für } r\geq 2\varrho.$$

Diese Formeln werden im folgenden nur für $m=0$ und $m=1$ benutzt.

3. Hilfssätze. *3.1.* Bekanntlich ([*13*], S. 183) ist

$$(3.1)\quad k_0(P,Q;-\lambda)=\frac{1}{4\pi r}e^{-r\sqrt{\lambda}} \quad \text{und} \quad \frac{dk_1(P,Q;-\lambda)}{dr}=-\frac{1}{8\pi}e^{-r\sqrt{\lambda}}.$$

Wegen (2.5), (2.6) und der für $r_{PQ}\leq R$ gültigen Ungleichungen

$$(3.2)\qquad |u_m(P,Q)| \quad \text{bzw.} \quad |\Delta u_m(P,Q)|\leq m_7 \qquad \text{für } m=0,1$$

ist also

$$|\Omega_0(P,Q;-\lambda)|\leq\frac{m_8}{r}e^{-r\sqrt{\lambda}} \qquad \text{für } 0\leq r\leq\varrho$$

und

$$|\Omega_0(P,Q;-\lambda)|\leq\frac{m_8}{r^3}e^{-\frac{r}{2}\sqrt{\lambda}} \qquad \text{für } \varrho\leq r\leq 2\varrho.$$

Folglich ist

$$(3.3)\qquad |\Omega_0(P,Q;-\lambda)|\leq\frac{m_9}{\varrho^2 r}e^{-\frac{r}{2}\sqrt{\lambda}} \quad \textit{und} \quad |\Gamma_0(P,Q;-\lambda)|\leq\frac{m_9}{r}e^{-r\sqrt{\lambda}}$$

für $0\leq r\leq 2\varrho$.

Ferner gilt:

$$(3.4)\qquad \Omega_1(P,Q;-\lambda) \quad \textit{und} \quad \frac{\partial}{\partial x^k}(\Omega_1(P,Q;-\lambda) \quad \textit{sind stetig auf} \quad V\times V.$$

Dabei ist entweder $(x^1,x^2,x^3)=P$ oder $=Q$. Die Behauptung (3.4) ist eine unmittelbare Folge der Formeln (2.6) und (3.1) sowie der Tatsache, daß $\frac{\partial}{\partial x^k}\Delta u_1(P,Q)$ auf $r_{PQ}\leq R$ stetig[4]) und $\eta(r)$ viermal stetig differenzierbar ist.

3.2. Ich beweise zunächst folgenden

Hilfssatz 1. *Es bezeichne* $G(P,Q;-\lambda)$ *die* GREENsche *Funktion von* $(\Delta-\lambda)u=0$, *und* $\Gamma_0(P,Q;-\lambda)$ *und* $\Omega_0(P,Q;-\lambda)$ *seien die in 2.1 bzw. 2.2 erklärten Funktionen. Dann hat die Integralgleichung*

$$(3.5)\quad \gamma(P)=\int_V\Gamma_0(\Pi,Q;-\lambda)\,\Omega_0(P,\Pi;-\lambda)\,dV_\Pi+\int_V\gamma(\Pi)\,\Omega_0(P,\Pi;-\lambda)\,dV_\Pi$$

[4]) Loc. cit.[1]).

für genügend große λ eine und nur eine Lösung, und es ist[5])

$$\gamma(P) = \gamma_0(P, Q; -\lambda) = \Gamma_0(P, Q; -\lambda) - G(P, Q; -\lambda).$$

Beweis. a) Es werde

$$\gamma_1(P, Q; -\lambda) = \Gamma_1(P, Q; -\lambda) - G(P, Q; -\lambda)$$

gesetzt, so daß wegen $(\Delta - \lambda) G = 0$

$$(\Delta - \lambda)\gamma_1(P, Q; -\lambda) = \Omega_1(P, Q; -\lambda)$$

ist. Wegen (3.4) ist auf Grund klassischer Sätze

$$\gamma_1(P, Q; -\lambda) = -\int_V G(P, \Pi; -\lambda)\,\Omega_1(\Pi, Q; -\lambda)\,dV_\Pi$$

und somit

$$\frac{\partial \gamma_1(P, Q; -\lambda)}{\partial x^k} = -\int_V G(P, \Pi; -\lambda)\,\frac{\partial}{\partial x^k}\,(\Omega_1(\Pi, Q; -\lambda)\,dV_\Pi$$

mit $(x^1, x^2, x^3) = Q$. Also:

(3.6) $\quad g^{ik}(Q)\,\dfrac{\partial \gamma_1(P, Q; -\lambda)}{\partial x^k}$ ist stetig auf $V \times V$.

Es werde

$$\gamma_0(P, Q; -\lambda) = \gamma_1(P, Q; -\lambda) - \eta(r_{PQ})\,u_1(P, Q)\,k_1(P, Q; -\lambda)$$

gesetzt. Wegen (3.6) und der zweiten Formel (3.1) ergibt sich:

(3.7) $\quad g^{ik}(Q)\,\dfrac{\partial \gamma_0(P, Q; -\lambda)}{\partial x^k}$ ist stetig auf $V \times V$.

Auf Grund der Definition von $\Gamma_m(P, Q; -\lambda)$ und $\gamma_1(P, Q; -\lambda)$ ist

$$\gamma_0(P, Q; -\lambda) = \Gamma_0(P, Q; -\lambda) - G(P, Q; -\lambda)$$

und somit

(3.8) $\quad (\Delta - \lambda)\,\gamma_0(P, Q; -\lambda) = \Omega_0(P, Q; -\lambda).$

b) Es bezeichne $\varkappa_Q^\varepsilon$ die geodätische Kugel um Q vom Durchmesser ε und $\bar{\varkappa}_Q^\varepsilon$ sei der Rand von $\varkappa_Q^\varepsilon$. Dann ist (Formel von GREEN)

$$\int_{V-\varkappa_Q^\varepsilon} \{\gamma_0(P, \Pi; -\lambda)\,(\Delta - \lambda)_\Pi\, G(\Pi, Q; -\lambda) - G(\Pi, Q; -\lambda)\,(\Delta - \lambda)_\Pi\,\gamma_0(P, \Pi; -\lambda)\}\,dV_\Pi$$

$$= \int_{\bar{\varkappa}_Q^\varepsilon} \left\{\gamma_0(P, \Pi; -\lambda)\,g^{ik}(\Pi)\,\frac{\partial G(\Pi, Q; -\lambda)}{\partial \xi^k} - G(\Pi, Q; -\lambda)\,g^{ik}(\Pi)\,\frac{\partial \gamma_0(P, \Pi; -\lambda)}{\partial \xi^k}\right\}\frac{\partial \xi^i}{\partial \nu}\,dF_\Pi.$$

[5]) Auf Grund des Prinzips der analytischen Fortsetzung und der Formel (4.2) läßt sich zeigen, daß (3.5) in der ganzen, längs der reellen negativen Achse aufgeschnittenen, komplexen λ-Ebene eine und nur eine Lösung hat. Diese Tatsache wird aber im folgenden nicht benutzt.

Dabei ist: $(\xi^1, \xi^2, \xi^3) = \Pi$, $dF_\Pi = \sqrt{g}\, dO$ mit dO = Flächenelement von $\vec{x}_Q^s$ und $\frac{\partial \xi^i}{\partial \nu}$ sind die Richtungskosinus der Flächennormalen auf $\vec{x}_Q^s$. Wegen (3.7) und (3.8) erhält man daraus nach Vollziehung des Grenzüberganges $\varepsilon \to 0$ die Darstellung

$$(3.9)\quad \begin{cases} \gamma_0(P, Q; -\lambda) = -\int\limits_V G(\Pi, Q; -\lambda)\, \Omega_0(P, \Pi; -\lambda)\, dV_\Pi \\ \qquad = -\int\limits_V \Gamma_0(\Pi, Q; -\lambda)\, \Omega_0(P, \Pi; -\lambda)\, dV_\Pi + \\ \qquad\quad + \int\limits_V \gamma_0(\Pi, Q; -\lambda)\, \Omega_0(P, \Pi; -\lambda)\, dV_\Pi. \end{cases}$$

c) Wegen (3.9) hat (3.5) mindestens eine Lösung nämlich $\gamma(P) = \gamma_0(P, Q; -\lambda)$. Daß für genügend große λ die Integralgleichung (3.5) außer $\gamma(P) = \gamma_0(P, Q; -\lambda)$ keine andere Lösung hat, sieht man folgendermaßen ein: Gäbe es zwei verschiedene Lösungen $v_1(\Pi)$ und $v_2(\Pi)$ von (3.5), so wäre

$$v_1(P) - v_2(P) = \int\limits_V \{v_1(\Pi) - v_2(\Pi)\}\, \Omega_0(P, \Pi; -\lambda)\, dV_\Pi.$$

Sei $\operatorname{Max} |v_1(P) - v_2(P)| = |v_1(P^*) - v_2(P^*)|$. Dann ist

$$|v_1(P^*) - v_2(P^*)| \leq |v_1(P^*) - v_2(P^*)| \int\limits_V |\Omega_0(P^*, \Pi; -\lambda)|\, dV_\Pi,$$

also

$$\int\limits_V |\Omega_0(P^*, \Pi; -\lambda)|\, dV_\Pi \geq 1.$$

Dies steht aber im Widerspruch mit der in 3.3 zu beweisenden Ungleichung (3.14), wonach $\int\limits_V |\Omega_0(P, \Pi; -\lambda)|\, dV_\Pi$ beliebig klein gemacht werden kann, wenn nur λ genügend groß gewählt wird.

3.3. Hilfssatz 2. *Es habe $\gamma_0(P, Q; -\lambda)$ dieselbe Bedeutung wie in Hilfssatz 1. Dann gibt es ein auf $V \times V$ erklärtes $R(P, Q; x)$ mit*

$$(3.10)\qquad |R(P, Q; x)| \leq m_1$$

und ein von P und Q unabhängiges $0 < 4\varrho_0 \leq R$, derart daß

$$\gamma_0(P, Q; -\lambda) = -\lambda \int\limits_0^\infty \frac{R(P, Q; x)}{(\lambda + x)^2}\, dx + O\left(e^{-\frac{\varrho_0}{2}\sqrt{\lambda}}\right), \qquad \lambda \to \infty$$

gleichmäßig in bezug auf $V \times V$ gilt.

Für $n = 1, 2, \ldots$ werde

$$\chi_n(P, \Pi_1, \ldots, \Pi_n, Q; -\lambda) = \Omega_0(P, \Pi_1; -\lambda)\, \Omega_0(\Pi_1, \Pi_2; -\lambda) \ldots \\ \ldots \Omega_0(\Pi_{n-1}, \Pi_n; -\lambda)\, \Gamma_0(\Pi_n, Q; -\lambda)$$

und

$$a_n(P, Q; -\lambda) = \overbrace{\int \cdots \int}^{n}_V \chi_n(P, \Pi_1, \ldots, \Pi_n, Q; -\lambda)\, dV_{\Pi_1} dV_{\Pi_2} \ldots dV_{\Pi_n}$$

gesetzt. Dann ist $-\sum\limits_{n=1}^{\infty} a_n(P, Q; -\lambda)$ die NEUMANNsche Reihe von (3.5).

Auf Grund des Hilfssatzes 1 ist für genügend große λ

$$\gamma_0(P, Q; -\lambda) = -\sum_{n=1}^{\infty} a_n(P, Q; -\lambda), \tag{3.11}$$

falls die rechtsstehende Reihe konvergiert. Ich beweise nicht nur, daß die erwähnte Reihe konvergiert, sondern ich beweise zugleich:

I. *Für jedes ϱ mit $0<2\varrho\leq R$ und jedes $\lambda\varrho^2\geq 1$ gibt es ein $r_n(P, Q; x)$ $(n=1, 2, \ldots)$ mit*

$$|r_n(P, Q; x)| \leq (m_{10}\varrho^2)^{n-\frac{1}{2}}, \tag{3.12}$$

so daß

$$\left| a_n(P, Q; -\lambda) - \lambda \int_0^{\infty} \frac{r_n(P, Q; x)}{(\lambda + x)^2}\, dx \right| \leq \frac{1}{\varrho^3}\left(\frac{m_{11}}{\varrho^2\lambda}\right)^{n-1} e^{-\frac{\varrho}{4}\sqrt{\lambda}} \tag{3.13}$$

ist.

Damit ist auch Hilfssatz 2 bewiesen. Denn wird

$$R(P, Q; x) = \sum_{n=1}^{\infty} r_n(P, Q; x)$$

und $m_{12}=\mathrm{Max}\{m_{10}, m_{11}\}$ gesetzt und werden die festen Zahlen $\varrho=4\varrho_0$ und λ^* so gewählt, daß

$$4m_{12}\varrho_0^2 \leq \frac{1}{2} \quad \text{und} \quad \frac{m_{12}}{4\varrho_0^3\lambda^*} \leq \frac{1}{2}$$

ist, so erhält man auf Grund von (3.12)

$$|R(P, Q; x)| \leq \sqrt{2}$$

und auf Grund von (3.11) und (3.13)

$$\left| \gamma_0(P, Q; -\lambda) + \lambda \int_0^{\infty} \frac{R(P, Q; x)}{(\lambda + x)^3}\, dx \right| \leq \frac{2}{\varrho_0^3} e^{-\frac{\varrho_0}{2}\sqrt{\lambda}},$$

wenn nur λ so groß gewählt wird, daß (3.11) gilt und zugleich $\lambda \geq \mathrm{Max}\left\{\lambda^*, \frac{1}{\varrho_0^2}\right\}$ ist.

Beweis von I. a) Es bezeichne $\varkappa_X^{\tau}$ die geodätische Kugel um X vom Durchmesser $0<\tau\leq R$ und ξ^1, ξ^2, ξ^3 sei ein Normalkoordinatensystem mit Zentrum X. Drücken wir die Koordinaten ξ^1, ξ^2, ξ^3 durch Polarkoordinaten r, α^1, α^2 aus, so daß $\xi^i = \Theta^i(\alpha^1, \alpha^2)r$, so ist die Transformationsdeterminante von der Gestalt $r^2\Psi(\Theta)$, und innerhalb $\varkappa_X^{\tau}$ gilt $|\Psi(\Theta)| \leq m_{13}$. Wegen (3.3) ist also

$$\begin{aligned} \int_V |\Omega_0(X, \Pi; -\lambda)|\, dV_{\Pi} &\leq \frac{m_9}{\varrho^2} \int_{\varkappa_X^{2\varrho}} e^{-r_{X\Pi}\sqrt{\lambda}/2} \frac{dV_{\Pi}}{r_{X\Pi}} \leq \\ &\leq \frac{m_{14}}{\varrho^2} \int_0^{2\pi}\int_0^{\pi} d\alpha^1\, d\alpha^2 \int_0^{2\varrho} e^{-\frac{r}{2}\sqrt{\lambda}}\, r\, dr \leq \frac{m_{15}}{\varrho^2\lambda}. \end{aligned} \tag{3.14}$$

Ebenso sieht man, daß

$$(3.15)\qquad \int\limits_V |\Omega_0(X,\Pi;-\lambda)|^2\,dV_\Pi \leq \frac{m_{16}}{\varrho^4\sqrt{\lambda}} \quad \text{und} \quad \int\limits_V |\Gamma_0(\Pi,Y;-\lambda)|^2\,dV_\Pi \leq \frac{m_{17}}{\sqrt{\lambda}}$$

ist. Infolgedessen ist (CAUCHY-SCHWARZsche Ungleichung)

$$(3.16)\qquad \int\limits_V |\Omega_0(X,\Pi;-\lambda)\,\Gamma_0(\Pi,Y;-\lambda)|\,dV_\Pi \leq \frac{m_{18}}{\varrho^3\sqrt{\lambda}}.$$

Wegen (3.3) ist schließlich

$$(3.17)\qquad \begin{cases} \displaystyle\int\limits_{V-\varkappa_Y^\varrho} |\Omega_0(X,\Pi;-\lambda)|\,dV_\Pi \leq \frac{m_9}{\varrho^3}\,e^{-\frac{\varrho}{2}\sqrt{\lambda}} \int\limits_{V-\varkappa_Y^\varrho} \frac{dV_\Pi}{r_{X\Pi}} \leq \frac{m_{19}}{\varrho^2}\,e^{-\frac{\varrho}{2}\sqrt{\lambda}} \leq \\ \qquad\qquad \leq \frac{1}{\varrho^3}\,\frac{m_{20}}{\sqrt{\lambda}}\,e^{-\frac{\varrho}{4}\sqrt{\lambda}} \end{cases}$$

und

$$\int\limits_{V-\varkappa_Y^\varrho} |\Omega_0(X,\Pi;-\lambda)|^2\,dV_\Pi \leq \frac{m_{21}}{\varrho^4}\,e^{-\varrho\sqrt{\lambda}}$$

und somit wegen (3.15)

$$(3.18)\qquad \begin{cases} \displaystyle\int\limits_{V-\varkappa_Y^\varrho} |\Omega_0(X,\Pi;-\lambda)\,\Gamma_0(\Pi,Y;-\lambda)|\,dV_\Pi \leq \\ \qquad \leq \left\{\int\limits_{V-\varkappa_Y^\varrho} |\Omega_0(X,\Pi;-\lambda)|^2\,dV_\Pi \int\limits_V |\Gamma_0(\Pi,Y;-\lambda)|^2\,dV_\Pi\right\}^{\frac{1}{2}} \leq \\ \qquad \leq \frac{m_{22}}{\varrho^3\sqrt{\lambda}}\,e^{-\frac{\varrho}{2}\sqrt{\lambda}} \leq \frac{1}{\varrho^3}\,\frac{m_{23}}{\varrho^2\lambda}\,e^{-\frac{\varrho}{4}\sqrt{\lambda}}. \end{cases}$$

b) Auf dieselbe Weise sieht man, daß

$$(3.19)\qquad \int\limits_{\varkappa_Y^\varrho} \frac{dV_\Pi}{r_{X\Pi}} \leq m_{24}\varrho^2, \qquad \int\limits_{\varkappa_Y^\varrho} \frac{|\eta(r_{\Pi Y})|}{r_{X\Pi}\,r_{\Pi Y}}\,dV_\Pi \leq m_{24}\varrho$$

ist.

c) Es ist

$$(3.20)\qquad \begin{cases} a_n(P,Q;-\lambda) = \left(\displaystyle\int\limits_{V-\varkappa_P^\varrho} + \int\limits_{\varkappa_P^\varrho}\right)\left(\int\limits_{V-\varkappa_{\Pi_1}^\varrho} + \int\limits_{\varkappa_{\Pi_1}^\varrho}\right)\cdots\left(\int\limits_{V-\varkappa_{\Pi_{n-1}}^\varrho} + \int\limits_{\varkappa_{\Pi_{n-1}}^\varrho}\right)\chi_n\,dV_{\Pi_1}\ldots dV_{\Pi_n} \\ \qquad = \displaystyle\int\limits_{V-\varkappa_P^\varrho}\left(\overbrace{\int\cdots\int}^{n=1}\right) + \cdots \qquad \cdots + \int\limits_{\varkappa_P^\varrho}\int\limits_{\varkappa_{\Pi_1}^\varrho}\cdots\int\limits_{\varkappa_{\Pi_{n-1}}^\varrho}. \end{cases}$$

Jeder Summand rechterhand, außer dem letzten enthält mindestens ein über $V-\varkappa_{\Pi_i}^\varrho$ $(i=0,1,\ldots,n-1$ mit $\Pi_0=P)$ erstrecktes Integral. Folglich

läßt sich jeder dieser Summanden mittels

$$\int_V |\Omega_0(\Pi_0, \Pi_1; -\lambda)|\, dV_{\Pi_1} \dots$$

$$\dots \int_V |\Omega_0(\Pi_{i-1}, \Pi_i; -\lambda)\, dV_{\Pi_i} \int_{V-\varkappa^{\varrho}_{\Pi_i}} |\Omega_0(\Pi_i, \Pi_{i+1}; -\lambda)|\, dV_{\Pi_{i+1}} \times$$

$$\times \int_V |\Omega_0(\Pi_{i+1}, \Pi_{i+2}; -\lambda)|\, dV_{\Pi_{i+2}} \dots \int_V |\Omega_0(\Pi_{n-1}, \Pi_n; -\lambda)\, \Gamma_0(\Pi_n, Q; -\lambda)|\, dV_{\Pi_n}$$

bzw.

$$\int_V |\Omega_0(\Pi_0, \Pi_1; -\lambda)|\, dV_{\Pi_1} \dots$$

$$\dots \int_V |\Omega_0(\Pi_{n-2}, \Pi_{n-1}; -\lambda)\, dV_{\Pi_{n-1}} \int_{V-\varkappa^{\varrho}_{\Pi_{n-1}}} |\Omega_0(\Pi_{n-1}, \Pi_n; -\lambda)\, \Gamma_0(\Pi_n, Q; -\lambda)|\, dV_{\Pi_n}$$

majorieren, je nachdem ob $0 \leq i \leq n-2$ oder $i = n-1$ ist.

Im ersten Falle ist, wegen (3.14), (3.16) und (3.17), dieser Ausdruck nicht größer als

$$\left(\frac{m_{15}}{\varrho^2 \lambda}\right)^{n-2} \frac{1}{\varrho^3} \frac{m_{20}}{\sqrt{\lambda}}\, e^{-\frac{\varrho}{4}\sqrt{\lambda}}\, \frac{m_{16}}{\varrho^3 \sqrt{\lambda}},$$

und im zweiten, wegen (3.14) und (3.18), nicht größer als

$$\left(\frac{m_{15}}{\varrho^2 \lambda}\right)^{n-1} \frac{1}{\varrho^3} \frac{m_{23}}{\varrho^2 \lambda}\, e^{-\frac{\varrho}{4}\sqrt{\lambda}}.$$

Wegen $\varrho^2 \lambda \geq 1$ gibt es also ein m_{25} derart, daß der betrachtete Ausdruck nicht größer als

$$\frac{1}{\varrho^3} \left(\frac{m_{25}}{\varrho^2 \lambda}\right)^{n-1} e^{-\frac{\varrho}{2}\sqrt{\lambda}}$$

ist. Da es in (3.20) derartiger Ausdrücke insgesamt $2^n - 1$ gibt, so ist

$$(3.21) \quad \left\{ \left| a_n(P, Q; -\lambda) - \int_{\varkappa^{\varrho}_P} \int_{\varkappa^{\varrho}_{\Pi_1}} \cdots \int_{\varkappa^{\varrho}_{\Pi_{n-1}}} \chi_n\, dV_{\Pi_1} \dots dV_{\Pi_n} \right| \leq \frac{2}{\varrho^3} \left(\frac{2 m_{25}}{\varrho^2 \lambda}\right)^{n-1} e^{-\frac{\varrho}{4}\sqrt{\lambda}} \leq \right.$$

$$\left. \leq \frac{1}{\varrho^3} \left(\frac{m_{11}}{\varrho^2 \lambda}\right)^{n-1} e^{-\frac{\varrho}{4}\sqrt{\lambda}}. \right.$$

d) Es werde

$$\chi_n(P, \Pi_1, \dots, \Pi_n, Q; -\lambda)$$

$$= \frac{1}{(4\pi)^{n+1}}\, \Delta u_0(P, \Pi_1) \dots \Delta u_0(\Pi_{n-1}, \Pi_n)\, \eta(r_{\Pi_n Q})\, u_0(\Pi_n, Q)\, \frac{e^{-(r_{P\Pi_1} + \dots + r_{\Pi_n Q})\sqrt{\lambda}}}{r_{P\Pi_1} \cdots r_{\Pi_n Q}}$$

$$= f_n(P, \Pi_1, \dots, \Pi_n, Q)\, \eta(r_{\Pi_n Q})\, \frac{e^{-(r_{P\Pi_1} + \dots + r_{\Pi_n Q})\sqrt{\lambda}}}{r_{P\Pi_1} \cdots r_{\Pi_n Q}}$$

gesetzt. Wegen (3.2) ist für $r_{\Pi_i \Pi_{i+1}} \in \varkappa^{\varrho}_{\Pi_i}$ $(i = 0, 1, \dots, n-1)$

$$(3.22) \qquad |f_n(P, \Pi_1, \dots, \Pi_n, Q)\, \eta(r_{\Pi_n Q})| \leq |\eta(r_{\Pi_n Q})| \left(\frac{m_4}{4\pi}\right)^{n+1}.$$

Wir setzen

$$v(r;x) = 1 - \frac{2}{\pi}\int_0^{r\sqrt{x}} \frac{\sin\xi}{\xi}\,d\xi,$$

so daß

$$e^{-r\sqrt{\lambda}} = \lambda\int_0^{\infty} \frac{v(r;x)}{(\lambda+x)^2}\,dx \tag{3.23}$$

und somit

$$\int_{\overset{\circ}{P}}\int_{\overset{\circ}{\Pi_1}}\cdots\int_{\overset{\circ}{\Pi_{n-1}}} \chi_n(P,\Pi_1,\ldots,\Pi_n,Q;-\lambda)\,dV_{\Pi_1}\ldots dV_{\Pi_n}$$

$$= \int_{\overset{\circ}{P}}\int_{\overset{\circ}{\Pi_1}}\cdots\int_{\overset{\circ}{\Pi_{n-1}}} \left\{\lambda\int_0^{\infty} v(r_{P\Pi_1}+\cdots+r_{\Pi_nQ};x)\frac{dx}{(\lambda+x)^2}\right\}\times$$

$$\times \frac{f_n(P,\Pi_1,\ldots,\Pi_n,Q)}{r_{P\Pi_1}\cdots r_{\Pi_nQ}}\eta(r_{\Pi_nQ})\,dV_{\Pi_1}\ldots dV_{\Pi_n}$$

$$= \lambda\int_0^{\infty}\frac{r_n(P,Q;x)}{(\lambda+x)^2}\,dx$$

mit

$$r_n(P,Q;x) = \int_{\overset{\circ}{P}}\int_{\overset{\circ}{\Pi_1}}\cdots$$

$$\cdots\int_{\overset{\circ}{\Pi_{n-1}}} v(r_{P\Pi_1}+\cdots+r_{\Pi_nQ};x)\frac{f_n(P,\Pi_1,\ldots,\Pi_n,Q)}{r_{P\Pi_1}\cdots r_{\Pi_nQ}}\eta(r_{\Pi_nQ})\,dV_{\Pi_1}\ldots dV_{\Pi_n}$$

ist. Offenbar ist $|v(r;x)| \le m_{26}$ und somit, wegen (3.22),

$$|r_n(P,Q;x)| \le m_{27}\left(\frac{m_7}{4\pi}\right)^{n+1}\int_{\overset{\circ}{P}}\frac{dV_{\Pi_1}}{r_{P\Pi_1}}\int_{\overset{\circ}{\Pi_1}}\frac{dV_{\Pi_2}}{r_{\Pi_1\Pi_2}}\cdots\int_{\overset{\circ}{\Pi_{n-1}}}\frac{|\eta(r_{\Pi_nQ})|}{r_{\Pi_{n-1}\Pi_n}r_{\Pi_nQ}}\,dV_{\Pi_n},$$

also wegen (3.19)

$$|r_n(P,Q;x)| \le m_{27}\left(\frac{m_7}{4\pi}\right)^{n+1}(m_{24}\varrho^2)^{n-1}m_{24}\varrho \le (m_{10}\varrho^2)^{n-\frac{1}{2}}\ ^{6}). \tag{3.24}$$

e) Aus (3.21) und (3.24) folgt die Behauptung.

4. Beweis des Satzes A. *4.1.* Dem Beweis des Satzes A schicke ich einige Hilfsbetrachtungen voraus.

Bekanntlich ist ([7], S. 452)

$$\sum_{n=1}^{\infty}\frac{|\Phi_n(P)\,\Phi_n(Q)|}{\lambda_n^2} \le m_{28} \tag{4.1}$$

---

[6]) In bezug auf die in d) benutzte Beweisanordnung verweise ich auf die Arbeit des Herrn R. BOJANIĆ [3], wo ein ähnlicher (im anderen Zusammenhang stehender) Beweisgedanke benutzt wird. Dort steht auch die Formel (3.23).

und für jedes $\lambda > 0$ und $\underline{\lambda} > 0$

$$(4.2)\quad \begin{cases} G(P,Q;-\lambda) - G(P,Q;-\underline{\lambda}) = -(\lambda-\underline{\lambda}) \sum_{n=0}^{\infty} \frac{\Phi_n(P)\,\Phi_n(Q)}{(\lambda_n+\underline{\lambda})(\lambda_n+\lambda)} \\ \qquad = \frac{1}{\lambda F} - \frac{1}{\underline{\lambda} F} - (\lambda-\underline{\lambda}) \sum_{n=1}^{\infty} \frac{\Phi_n(P)\,\Phi_n(Q)}{(\lambda_n+\underline{\lambda})(\lambda_n+\lambda)}, \end{cases}$$

da $\lambda_0 = 0$ und $\Phi_0(P) = \frac{1}{\sqrt{F}}$ ist. Folglich existiert auf $V \times V$

$$a(P,Q) = \lim_{\underline{\lambda}\to 0} \left\{ G(P,Q;-\underline{\lambda}) - \frac{1}{\underline{\lambda} F} \right\},$$

so daß

$$(4.3)\quad \begin{cases} \lambda \sum_{n=1}^{\infty} \frac{\Phi_n(P)\,\Phi_n(Q)}{\lambda_n(\lambda_n+\lambda)} = a(P,Q) - G(P,Q;-\lambda) + \frac{1}{\lambda F} \\ \qquad = a(P,Q) - \Gamma_0(P,Q;-\lambda) + \frac{1}{\lambda F} + \gamma_0(P,Q;-\lambda) \end{cases}$$

ist. Wegen (4.1) und (4.3) existiert $\lim_{Q\to P} \{a(P,Q) - \Gamma_0(P,Q;-\lambda)\}$. Infolgedessen und wegen

$$a(P,Q) - \Gamma_0(P,Q;-\lambda) = a(P,Q) - \frac{u_0(P,Q)}{4\pi r_{PQ}} e^{-r_{PQ}\sqrt{\lambda}}$$

$$= a(P,Q) - \frac{u_0(P,Q)}{4\pi r_{PQ}} + \frac{u_0(P,Q)}{4\pi}\sqrt{\lambda} - \frac{u_0(P,Q)}{4\pi r_{PQ}} \sum_{n=2}^{\infty} (-1)^n \frac{(r_{PQ}\sqrt{\lambda})^n}{n!}$$

sowie $u_0(P,P) = 1$ gibt es ein $w(P)$ derart, daß

$$(4.4)\qquad \lim_{Q\to P} \{a(P,Q) - \Gamma_0(P,Q;-\lambda)\} = \frac{\sqrt{\lambda}}{4\pi} + w(P).$$

Dabei ist offenbar $w(P)$ stetig auf V und somit

$$(4.5)\qquad |w(P)| \le m_2,$$

da V geschlossen ist.

Aus (4.3) und (4.4) folgt schließlich

$$(4.6)\qquad \lambda \sum_{n=1}^{\infty} \frac{\Phi_n^2(P)}{\lambda_n(\lambda_n+\lambda)} = \frac{\sqrt{\lambda}}{4\pi} + w(P) + \frac{1}{\lambda F} + \gamma_0(P,P;-\lambda).$$

Es bezeichne $R(P,Q;x)$ die in Hilfssatz 2 erklärte Funktion und es werde $R(P;x) = R(P,P;x)$ gesetzt. Durch die hier gegebenen Definitionen von $w(P)$ und $R(P;x)$ ist das $A(P;x)$ des Satzes A, d.h.

$$A(P;x) = \sum_{\lambda_1 \le \lambda_n \le x} \frac{\Phi_n^2(P)}{\lambda_n} - \frac{\sqrt{x}}{2\pi^2} - w(P) + R(P;x)$$

vollständig erklärt, und wegen (3.10) und (4.5) genügen $w(P)$ und $R(P;x)$ den Bedingungen (1.3).

Nun beweise ich: *Wird*

$$A^*(P;x)=\int_0^x A(P;t)\,dt-\frac{1}{F}$$

und

$$f(P;-\lambda)=\lambda\int_0^\infty \frac{A^*(P;x)}{(\lambda+x)^3}\,dx$$

gesetzt, so ist

$$A^*(P;x)=O(x^{\frac{1}{2}}),\qquad x\to\infty \tag{4.7}$$

und

$$|f(P;-\lambda)|\leq m_{30}\frac{1+\lambda}{\lambda}e^{-\frac{\varrho_0}{2}\sqrt{\lambda}}. \tag{4.8}$$

Beweis. a) Es habe $\Theta_1(P;x)$ dieselbe Bedeutung wie in 1.5, d.h. es sei $\Theta_1(P;x)=\sum_{\lambda_1\leq\lambda_n\leq x}\Phi_n^2(P)/\lambda_n$, so daß die linke Seite von (4.6) durch

$$\lambda\int_0^\infty d\{\Theta_1(P;x)\}/(\lambda+x)$$

dargestellt werden kann. Wegen $\Theta_1(P;0)=0$ erhält man daraus durch gliedweise Integration die Darstellung

$$\lambda\sum_{n=1}^\infty\frac{\Phi_n^2(P)}{\lambda_n(\lambda_n+\lambda)}=\lambda\int_0^\infty\frac{\Theta_1(P;x)}{(\lambda+x)^2}\,dx. \tag{4.9}$$

b) Wegen der ersten Ungleichung (1.3) ist

$$\left|\lambda\int_0^\infty\frac{R(P;x)}{(\lambda+x)^2}\,dx\right|\leq m_1$$

und somit auf Grund des Hilfssatzes 2

$$\gamma_0(P,P;-\lambda)=O(1)\quad(\lambda\to\infty). \tag{4.10}$$

Aus (4.6), (4.9) und (4.10) folgt

$$\lambda\int_0^\infty\frac{\Theta_1(P;x)}{(\lambda+x)^2}\,dx=O(\sqrt{\lambda})\quad(\lambda\to\infty).$$

Da $\Theta_1(P;x)$ nicht abnimmt, ist

$$\Theta_1(P;\lambda)=\Theta_1(P;\lambda)\,2\lambda\int_\lambda^\infty\frac{dx}{(\lambda+x)^2}\leq 2\lambda\int_0^\infty\frac{\Theta_1(P;x)}{(\lambda+x)}\,dx=O(\sqrt{\lambda})$$

und somit

$$A(P;x)=O(\sqrt{x}), \tag{4.11}$$

woraus sich (4.7) unmittelbar ergibt.

c) Wegen $\lambda\int_0^\infty \sqrt{\lambda}\,dx/(\lambda+x)^2=\pi\sqrt{\lambda}/2$ und (4.9) ist

$$(4.12)\quad \lambda\int_0^\infty \frac{A(P;x)}{(\lambda+x)^2}\,dx=\lambda\sum_{\lambda_1\le\lambda_n\le x}\frac{\Phi_n^2(P)}{\lambda_n}-\frac{\sqrt{\lambda}}{4\pi}-w(P)+\lambda\int_0^\infty\frac{R(P;x)}{(\lambda+x)^2}\,dx.$$

Wird das linksstehende Integral partiell integriert und $2\lambda\int_0^\infty dx/(\lambda+x)^3=1/\lambda$ beachtet, so sieht man, daß die linke Seite von (4.12) gleich $2f(P;-\lambda)+1/\lambda F$ ist. Infolgedessen und nach Formel (4.6) erhält man die Beziehung

$$2f(P;-\lambda)=\gamma_0(P,P;-\lambda)+\lambda\int_0^\infty\frac{R(P;x)}{(\lambda+x)^2}\,dx.$$

d) Daraus folgt (auf Grund des Hilfssatzes 2) $f(P;-\lambda)=O\left(e^{-\frac{\varrho_0}{2}\sqrt{\lambda}}\right)$, und zwar gleichmäßig in bezug auf P.

e) Wegen (4.1), (4.6) und (4.5) ist

$$|\gamma_0(P,P;-\lambda)|\le\frac{\sqrt{\lambda}}{4\pi}+m_2+\frac{1}{\lambda F}+\lambda m_{28},\ \text{d.h. } \gamma_0(P,P;-\lambda)=O\left(\frac{1}{\lambda}\right),\ \lambda\to 0,$$

und zwar gleichmäßig in bezug auf P.

f) Aus d) und e) folgt (4.8).

4.2. Nun gehe ich zum letzten Schritt des Beweises über. Den Ausgangspunkt bilden die Ungleichungen (4.7) und (4.8) und die bekannte Formel

$$\frac{e^{-|s|\sqrt{u}}}{\sqrt{u}}=\int_0^\infty\frac{\cos s\sqrt{\lambda}}{\sqrt{\lambda}}\,\frac{d\lambda}{\lambda+u}\qquad (s\text{ reell}).$$

Durch zweimalige gliedweise Integration erhält man daraus die Formel

$$\frac{e^{-|s|\sqrt{u}}}{\sqrt{u}}=\int_0^\infty\frac{h(s;\lambda)}{(\lambda+u)^3}\,d\lambda$$

mit

$$(4.13)\qquad h(s;\lambda)=\frac{2}{\pi}\int_0^\lambda dt_1\int_0^{t_1}\frac{\cos s\sqrt{t_2}}{\sqrt{t_2}}\,dt_2.$$

Wegen (4.8) und der unmittelbar ersichtlichen Ungleichung $|h(s;\lambda)|\le\frac{8}{3\pi}\lambda^{\frac{3}{2}}$ existiert das über $(0,\infty)$ erstreckte Integral von $f(P;-\lambda)\,h(s;\lambda)/\lambda$. Im derart gebildeten Doppelintegral läßt sich wegen (4.7) die Reihenfolge der Integration vertauschen (Satz von FUBINI) und man erhält für $s>0$

$$(4.14)\quad \begin{cases}\displaystyle\int_0^\infty f(P;-\lambda)\,h(s;\lambda)\,\frac{d\lambda}{\lambda}=\int_0^\infty A^*(P;u)\,du\int_0^\infty\frac{h(s;\lambda)}{(\lambda+u)^3}\,d\lambda\\ \displaystyle=\int_0^\infty e^{-s\sqrt{u}}\,A^*(P;u)\,\frac{du}{\sqrt{u}}=\varphi^*(P;s).\end{cases}$$

Ich beweise zunächst: $\varphi^*(P;s)$ *ist in der rechten Halbebene und im Kreis* $|s|<\frac{\varrho_0}{2}$ *regulär und für* $|s|\leq\frac{\varrho_0}{3}$ *gleichmäßig beschränkt in bezug auf* s *und* P.

Beweis. a) Wegen (4.7) und (4.8) konvergieren für positive s sowohl die linke Seite von (4.14) als auch das an zweiter Stelle rechts stehende LAPLACE-Integral. Bezüglich des ersten Teiles der Behauptung haben wir also nur noch zu zeigen, daß die linke Seite von (4.14) absolut konvergiert, falls $|s|<\frac{\varrho_0}{2}$. Dies ist aber eine unmittelbare Folge von (4.8) und der für jedes komplexwertige s gültigen Ungleichung

$$|h(s;\lambda)|\leq\frac{2}{\pi}e^{|s|\sqrt{\lambda}}\int_0^{\lambda}dt_1\int_0^{t_1}\frac{dt_2}{\sqrt{t_2}}\leq\frac{8}{3\pi}\lambda^{\frac{1}{2}}e^{|s|\sqrt{\lambda}}. \tag{4.15}$$

b) Wegen (4.8) und (4.14) ist für $|s|\leq\frac{\varrho_0}{3}$,

$$|\varphi^*(P;s)|\leq\frac{8m_{30}}{3\pi}\int_0^{\infty}e^{-\left(\frac{\varrho_0}{2}-|s|\right)\sqrt{\lambda}}\frac{1+\lambda}{\sqrt{\lambda}}d\lambda\leq m_{31}\int_0^{\infty}e^{-\frac{\varrho_0}{6}\sqrt{\lambda}}\frac{1+\lambda}{\sqrt{\lambda}}d\lambda$$

$$\leq m_{32}\left(\frac{1}{\varrho_0}+\frac{1}{\varrho_0^3}\right),$$

womit auch der zweite Teil der Behauptung bewiesen ist.

Wegen

$$\int_0^X e^{-sv}d\{A^*(P;v^2)\}=A^*(P;X^2)\,e^{-sX}-A^*(P;0)+s\int_0^X e^{-sv}A^*(P;v^2)\,dv$$

und (4.7) ist für $\Re(s)>0$

$$\int_0^{\infty}e^{-sv}d\{A^*(P;v^2)\}=-A^*(P;0)+s\int_0^{\infty}e^{-sv}A^*(P;v^2)\,dv.$$

Denn wegen (4.14) existiert das rechtsstehende Integral und ist gleich $\frac{s}{2}\varphi^*(P;s)$. Also ist

$$\frac{S}{2}\varphi^*(P;s)=A^*(P;0)+\int_0^{\infty}e^{-sv}d\{A^*(P;v^2)\}.$$

Wegen $A^*(P;0)=-\frac{1}{F}$ und $d\{A^*(P;v^2)\}=2vA(P;v^2)\,dv$ erhält man schließlich

$$\frac{S}{2}\varphi^*(P;s)=-\frac{1}{F}+2\int_0^{\infty}e^{-sv}v\,A(P;v^2)\,dv.$$

Infolgedessen ist

$$\varphi^{**}(P;s)=\int_0^{\infty}e^{-sv}v\,A(P;v^2)\,dv \tag{4.16}$$

in der rechten Halbebene einschließlich $|s|<\frac{\varrho_0}{2}$ regulär und für $|s|\leq\frac{\varrho_0}{3}$ gleichmäßig beschränkt in bezug auf s und P.

Daraus folgt: *Für alle in der rechten Halbebene oder im Kreis* $|s| < \frac{\varrho_0}{2}$ *liegenden* s *existiert*

$$\varphi(P;s) = \int_S^\infty \varphi^{**}(P;Z)\,dZ \tag{4.17}$$

und ist daselbst regulär und für $|s| \leq \frac{\varrho_0}{3}$ *gleichmäßig beschränkt in bezug auf* s *und* P.

Wegen (4.16) und (4.17) ist für $\Re(s) > 0$

$$\varphi(P;s) = \int_0^\infty e^{-sv} A(P;v^2)\,dv,$$

womit der Beweis des Satzes A beendet ist[7]).

5. Das Gegenbeispiel. *5.1.* V sei durch $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$ definiert und Δ durch Polarkoordinaten $x_1 = \sin x^1, \ldots, x_4 = \sin x^1 \sin x^2 \sin x^3$ ausgedrückt. Dann ist $F = 2\pi^2$ und $n(n+2)$ sind die Eigenwerte von $(\Delta + \lambda)u = 0$. Jeder Eigenwert ist $(n+1)^2$-fach besetzt. $w_n^i(P)$ $\left(i = 1, 2, \ldots, (n+1)^2\right)$ seien die dem Eigenwert $n(n+2)$ entsprechenden linear unabhängige Eigenfunktion. Dann ist (vgl. [6] und [11])

$$\int_V w_n^i(\Pi)\, w_n^j(\Pi)\,dV_n = 0 \quad \text{bzw. } 1$$

je nachdem ob $i \neq j$ oder $i = 1$ ist, und

$$\sum_{i=1}^{i=(n+1)^2} w_n^i(P)\, w_n^i(Q) = C_n^3(\cos\vartheta)\,\frac{1}{2\pi^2}. \tag{5.1}$$

Dabei ist $\cos\vartheta = \sum_{i=1}^{4} x_i y_i$ [8]) und die $C_n^3(\cos\vartheta)$ sind mittels

$$\sum_{n=0}^{\infty} C_n^3(\cos\vartheta)\, z^n = \frac{1 - z^2}{(1 - 2z\cos\vartheta + z^2)^2}$$

erklärt, so daß

$$C_n^3(1) = (n+1)^2 \tag{5.2}$$

ist.

5.2. Wegen (5.1) und (5.2) ist für ganzzahlige x

$$\Theta\big(P; x(x+2)\big) = \frac{1}{2\pi^2} \sum_{n=0}^{n=x} (n+1)^2 = \frac{1}{12\pi^2}\,(2x+3)\,(x+2)\,(x+1),$$

also

$$\frac{\Theta\big(P; x(x+2)\big) - \frac{1}{6\pi^2}\,(x+2)^{\frac{3}{2}}\, x^{\frac{3}{2}}}{(x+2)\,x} = \frac{x}{6\pi^2}\left[\left(1 + \frac{3}{2x}\right)\left(1 + \frac{1}{x}\right) - \left(1 + \frac{2}{x}\right)^{\frac{1}{2}}\right]$$

[7]) Die in 4.2 benutzte Beweismethode ist im Grunde dieselbe, welche ich in [1] verwendet habe. Diesbezüglich sei auch auf die Arbeit des Herrn V. VUČKOVIĆ [12] verwiesen.

[8]) $x_1, \ldots, x_4$ bzw. $y_1, \ldots, y_4$ sind die Koordinaten von P bzw. Q.

und somit

$$\lim_{x\to\infty}\frac{\Theta(P;x(x+2))-\frac{1}{6\pi^2}(x+2)^{\frac{1}{2}}x^{\frac{1}{2}}}{(x+2)x}=\frac{1}{4\pi^2}.$$

Wegen $\sum_{\lambda_n\leq(x+2)x}1=\sum_{n=0}^{n=x}(n+1)^2$ ist schließlich

$$\lim_{x\to\infty}\frac{\sum_{\lambda_n\leq(x+2)x}1-\frac{1}{3}(x+2)^{\frac{1}{2}}x^{\frac{1}{2}}}{(x+2)x}=\frac{3}{4}.$$

Literatur

[1] AVAKUMOVIĆ, V. G.: Bemerkung über einen Satz des Herrn T. CARLEMAN. Math. Z. **53**, 53—58 (1950). — [2] AVAKUMOVIĆ, V. G.: Über die Eigenfunktionen der Schwingungsgleichung. Publ. Math. Inst. Math. Acad. Serbe Sci. **4**, 95—96 (1952). — [3] BOJANIĆ, R.: Propriétés asymptotiques des solutions des équations différentielles lineaires. Recueil Trav. Acad. Serbe Sci. **35**, Inst. Math. **3**, 213—254 (1953). — [4] CARLEMAN, T.: Propriétés asymptotiques des fonctions fondamentales des membranes vibrantes. C. R. Skand. Math. Kongress **1934**, 34—44. — [5] COURANT-HILBERT: Methoden der Mathematischen Physik, Bd. I. 1931. — [6] HEINE: Kugelfunktionen. 1878. — [7] GOURSAT: Cours d'Analyse, Bd. III. 1942. — [8] INGHAM, A. E.: On Wiener's method in Tauberian theorems. Proc. London Math. Soc. **38**, 458—480 (1935). — [9] KARAMATA, J.: Über einen Satz von H. HEILBRON und E. LANDAU. Publ. Math. Univ. Belgrad **5**, 28—38 (1936). — [9a] LEVITAN, B. M.: Über Entwickelungen nach den Eigenfunktionen des LAPLACEschen Operators. (Russisch.) Dokl. Akad. Nauk SSSR. **90**, 133—135 (1953). — [10] MINAKSHISUNDARAM, S., and A. PLEIJEL: Some properties of eigenfunctions of the LAPLACE-operator on RIEMANNian manifolds. Canadian J. Math. **1**, 242—256 (1949). — [11] MINAKSHISUNDARAM, S.: Zeta-functions on the Sphere. J. Indian Math. **13**, 41—48 (1949). — [12] VUČKOVIĆ, V.: Die STIELTJES-Transformation, die mit Geschwindigkeit der Exponentialfunktion unendlich klein wird. Recueil Trav. Acad. Serbe Sci. **35**, Inst. Math. **3**, 255—288 (1953). — [13] WATSON, G. N.: A Treatise on the Theory of BESSEL Functions. 1944. — [14] WEYL, H.: Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen. Math. Ann. **71**, 441—471 (1911).

Institut Mathematique de l'Academie Serbe des Sciences Belgrade (Yougoslavie)

(Eingegangen am 1. Dezember 1955)

FOULKES, H. O.
Math. Zeitschr. Bd. 65, S. 345—352 (1956)

Theorems of Kakeya and Pólya on power-sums

By
H. O. FOULKES

1. Introduction

In 1925, KAKEYA [1] proved that any symmetric function of n indeterminates x_i, $i=1, \ldots, n$ could be expressed as a rational function of a fundamental set

(I) $$s_{\nu_1}, s_{\nu_2}, \ldots, s_{\nu_n},$$

where $s_k = \sum_{i=1}^{n} x_i^k$ and $\nu_1, \nu_2, \ldots, \nu_n$ are the first n positive integers which do not belong to an additive sequence, or semigroup, of positive integers. By choosing the semigroup in different ways various special cases arise. These have appeared as theorems of LAGUERRE [2], BORCHARDT [3], VAHLEN [4] and LUDWIG [5]. KAKEYA's proof however gives no means of expressing any given symmetric function, even of the simplest kind such as s_k, $k \neq \nu_i$ or $a_r = \sum x_1 x_2 \ldots x_r$, $r \leq n$, in the required rational form. This problem has been attacked by some of the above writers and also by NAKAMURA [6], PÓLYA [7] and PERRON [8], [9] but with only partial success.

In this paper I use Schur functions to give a short proof of KAKEYA's theorem which has the additional merit that it may be used to derive simple explicit expressions for the functions s_k and a_r. Furthermore I give a new proof of a theorem of PÓLYA which in the same way gives a precision to the theorem which was lacking in the original proof.

2. Preliminary results on S-functions

Let $(\lambda) = (\lambda_1, \lambda_2, \ldots, \lambda_r)$ be any partition of a positive integer m, with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$. Then the SCHUR function $\{\lambda\}$ is defined as [10]

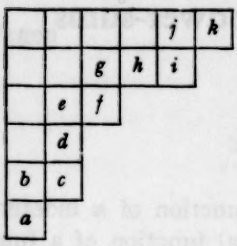
$$\{\lambda\} = \frac{1}{m!} \sum_{\varrho} h_\varrho \chi_\varrho^\lambda S_\varrho,$$

where $(\varrho) = (1^{\varrho_1} 2^{\varrho_2} \ldots)$ is a partition of m defining a conjugate class of order h_ϱ of the symmetric group of order $m!$, χ_ϱ^λ is the characteristic of the class (ϱ) in the irreducible (λ) representation of the group, and $S_\varrho = s_1^{\varrho_1} s_2^{\varrho_2} \ldots$. We shall also need the differential operator [11]

$$D_\lambda = \sum \frac{\chi_\varrho^\lambda}{\varrho_1! \varrho_2! \cdots} \frac{\partial^{\varrho_1+\varrho_2+\cdots}}{\partial s_1^{\varrho_1} \partial s_2^{\varrho_2} \ldots},$$

which is obtained from $\{\lambda\}$ by replacing every s_i^a by $i^a \frac{\partial^a}{\partial s_i^a}$. I have shown [11] that $D_\lambda \{\mu\} = \sum g_{\nu\lambda\mu} \{\nu\}$, where $\{\nu\}\{\lambda\} = \sum g_{\nu\lambda\mu} \{\mu\}$.

Partitions will be represented by diagrams of squares in the usual way. If in a diagram we can start at the bottom of any column and, working to the right or upwards or both in any succession, remove a continuous strip of α squares from the edge of the diagram so as to finish at the right hand end of a row, then we say that the diagram has a removable α-strip [12]. Thus when $(\lambda) = (6532^21)$, we have two removable 6-strips, $abcdef$ and $fghijk$. The removable 4-strips are $abcd$, $cdef$, $fghi$ and $hijk$. There are no removable 5, 8 or 10-strips. In effect, a strip is removable when its deletion leaves a proper partition diagram, or possibly no diagram at all.



We need three simple well known results from S-function theory.

Lemma 1. *An S-function $\{\lambda_1, \lambda_2, \ldots, \lambda_r\}$ of n indeterminates vanishes identically if $r > n$,* ([10], p. 91).

Lemma 2. *If the (λ)-diagram has no removable α-strip, then χ_ϱ^λ is zero for any S_ϱ which contains a non-zero power of s_α* ([10], p. 68–71; [12]).

Lemma 3. *If the (λ)-diagram has a removable α-strip, then χ_ϱ^λ, where S_ϱ contains a non-zero power of s_α, is non-zero for at least one ϱ.*

This follows since if a (μ)-diagram is obtained by removing the α-strip, then not all the characteristics of the irreducible (μ) representation of the symmetric group of order $(m-\alpha)!$ can be zero.

3. KAKEYA'S Theoren.

An additive sequence A will be determined by a set of basis integers, and will consist of multiples of the greatest common divisor d of this set. It will contain all the multiples of d greater than some fixed value, and so if $d = 1$, A will contain all but a finite number of the positive integers and there will be a greatest value N of n. If $d \neq 1$ there is no restriction on n. In the following, when a value of n is chosen it should be understood that it is a possible one for the particular A, that is, it cannot exceed N.

KAKEYA'S Theorem. *If $\nu_1, \nu_2, \ldots, \nu_n$ are the first n positive integers which do not belong to a given additive sequence, then for n indeterminates every s_r can be expressed as a rational function of $s_{\nu_1}, s_{\nu_2}, \ldots, s_{\nu_n}$.*

Proof. Let α be the first member of the additive sequence A. Construct a partition diagram, starting at its lowest point, by erecting a column of α squares. Then start a row of squares to the right using the top square of the α-column as starting point. As soon as the total number of squares in the column and the row is a number x, not in A, start erecting another column with the x-square as its lowest point. Continue this column until the total number of squares in the two columns and the single row is a number

β of A. Now turn again from the vertical to the horizontal and continue the row to the right until the total number of squares in the two columns and the two rows is a number y not in A. Continuing in this way, changing from column to row and row to column as dictated by the numbers $\alpha, \beta, \ldots \in A$ and $x, y, \ldots \notin A$, we obtain the edge of a partition diagram which is entirely determined by A. If $d=1$ there will be a finite number of rows in the diagram but the top row will extend indefinitely to the right. If $d \neq 1$ the number of rows will be infinite.

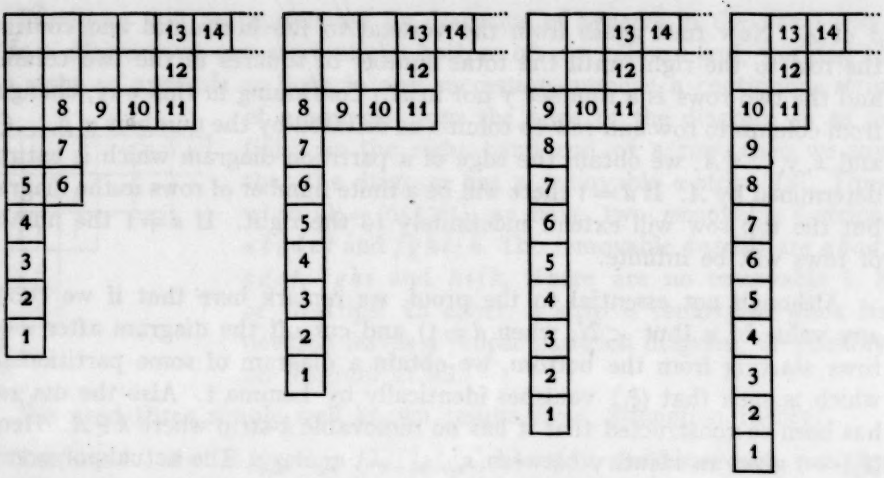
Although not essential to the proof, we remark here that if we choose any value of n (but $<N$, when $d=1$) and cut off the diagram after $n+1$ rows starting from the bottom, we obtain a diagram of some partition (ξ_1) which is such that $\{\xi_1\}$ vanishes identically by Lemma 1. Also the diagram has been so constructed that it has no removable k-strip where $k \in A$. Hence $\{\xi_1\}=0$ gives an identity between $s_{\nu_1}, s_{\nu_2}, \ldots, s_{\nu_n}, s_{\nu_{n+1}}$. The actual polynomial can be written down from the group character tables or by using some of the known rules for finding particular characteristics.

Continuing with the proof, delete the second column of the full partition diagram and move the complete first column one place to the right to close the gap. Then add one square to the bottom of the first column. It is now possible to remove an α-strip from the resulting diagram, but only one such strip can be removed, and there is no other removable k-strip, where $k \in A$. Take $n+1$ rows of this diagram thus obtaining the diagram of some partition (ξ_2). By Lemma 1, $\{\xi_2\}=0$, and so by Lemma 2 we have a polynomial identity between $s_\alpha, s_{\nu_1}, s_{\nu_2}, \ldots, s_{\nu_n}$ in which s_α occurs to the first degree only, and occurs with a non-zero coefficient by Lemma 3. Thus s_α is expressed as a rational function of (I).

Next delete the second column of the last full diagram, move the first column one place to the right and add one square to the bottom of the first column, exactly as at the preceding stage. It will now be possible to remove one, or more, α-strips from the new diagram, but if α' is the second member of A, there will be one and only one removable α'-strip and there will be no other removable k-strips, where $k \in A$. Take $n+1$ rows of this diagram, giving a partition (ξ_3). Then $\{\xi_3\}=0$ gives a polynomial identity between $s_\alpha, s_{\alpha'}, s_{\nu_1}, \ldots, s_{\nu_n}$ in which $s_{\alpha'}$ appears with non-zero coefficient and only to the first degree. Hence $s_{\alpha'}$ is expressed as a rational function of (I).

When $d \neq 1$, we can proceed in precisely the same way to express every s_k, $k \in A$, as a rational function of (I). When $d=1$, the various identities $\{\xi_i\}=0$ will at some stage become $\{1^{N+1}\}=0$. This expresses s_{N+1} in the required form, and similarly every s_{N+l}, $l=2, 3, \ldots$, is obtained from $\{1^{N+l}\}=0$. This completes the proof.

EXAMPLE 1. Take 5, 8, 9 as basis so that A is 5, 8, 9, 10, 13, 14, ... and the possible values for $\nu_1, \nu_2, \ldots, \nu_n$ will be taken from 1, 2, 3, 4, 6, 7, 11, 12. Here $d=1$, $N=8$ and the first four diagrams are



When $n=6$, $\{5^2 1^4\}=0$ gives s_{11} as a rational function of $s_1, s_2, s_3, s_4, s_6, s_7$, and $\{5^2 2^2 1^4\}=0$ would give s_{12} as a rational function of $s_1, s_2, s_3, s_4, s_6, s_7, s_{11}$. These can however be obtained from S-functions of lower weight, for $\{1^7\}=0$ gives

$$s_5=(s_1^7-21s_1^5s_2+70s_1^4s_3-210s_1^3s_4+105s_1^3s_2^2-420s_1^2s_2s_3-840s_1s_6$$
$$+630s_1s_2s_4-105s_1s_2^3+280s_1s_3^2+210s_2^2s_3-420s_3s_4+720s_7)/504(s_2-s_1^2),$$

$\{1^8\}=0$ gives s_8, and so on for every further s_r.

When $n=7$, $\{5^2 2^2 1^4\}=0$ gives s_{12} in the required form as above, and $\{41^7\}=0$ gives s_5 as a rational function of $s_1, s_2, s_3, s_4, s_6, s_7, s_{11}$. For s_r, $r\geq 8$, we use $\{1^r\}=0$.

When $n=8$, we use $\{4^2 1^7\}=0$, $\{31^8\}=0$, $\{1^r\}=0$ to give respectively s_5, s_8, s_r, where $r=9, 10, 13, \ldots$, as rational functions of $s_1, s_2, s_3, s_4, s_6, s_7, s_{11}, s_{12}$.

EXAMPLE 2. If the additive sequence consists merely of all the multiples of an integer α, we have the case considered by VAHLEN [4]. The successive partition diagrams obtained in the above proof are those of $(\ldots 3^{\alpha-1}2^{\alpha-1}1^{\alpha-1})$, $(\ldots 3^{\alpha-1}2^{\alpha-1}1^{2\alpha-1})$, $(\ldots 3^{\alpha-1}2^{\alpha-1}1^{3\alpha-1})$ and so on. When $\alpha=2$, the initial partition diagram is $(\ldots 4321)$ and the first identity obtained is $\{n+1, n, \ldots, 321\}=0$. For $n=1, 2, 3$ this gives $s_1^3-s_3=0$,

$$s_1^6-5s_1^3s_3+9s_1s_5-5s_3^2=0,$$

$$s_1^{10}-15s_1^7s_3+63s_1^5s_5-225s_1^3s_7+315s_1^2s_3s_5-175s_1s_3^3+225s_3s_7-189s_5^2=0.$$

These expressions appeared, but not perhaps in such a natural and predictable fashion, in the papers of PÓLYA [7] and PERRON [8].

Taking $n=3$, the last of these identities gives s_7 as a rational function of s_1, s_3, s_5 and from $\{21^3\}=0$ we have

$$s_2=\frac{s_1^5+5s_1^2s_3-6s_5}{5(s_1^3-s_3)}.$$

We deal with $s_4, s_6, s_8, s_9, \ldots$ by using $\{1^r\}=0$, $r=4, 6, 8, 9, \ldots$.

4. Expressions for the elementary symmetric functions a_r as rational functions of a fundamental set

THEOREM. *If (λ) is the partition of n rows determined by a given additive sequence and a given n, then the expression for a_r as a rational function of the fundamental set is*

$$a_r = \frac{D_{1^{n-r}}\{\lambda\}}{D_{1^n}\{\lambda\}}.$$

PROOF. For n indeterminates, $\{\lambda\}$ will not vanish identically and is a polynomial in $s_{\nu_1}, s_{\nu_2}, \ldots, s_{\nu_n}$. Also $\{\lambda_1-1, \lambda_2-1, \ldots, \lambda_n-1\}$ is a non-vanishing polynomial in the same power sums since it is obtained from $\{\lambda\}$ by differentiating in accordance with $D_{1^n}\{\lambda\} = \{\lambda_1-1, \lambda_2-1, \ldots, \lambda_n-1\}$.

Since we have n indeterminates only, the product $a_r\{\lambda_1-1, \lambda_2-1, \ldots, \lambda_n-1\}$ will contain only those S-functions whose partitions can be obtained by taking from the edge of the (λ)-diagram $n-r$ squares no two of which are in the same row. But these are precisely the S-functions given by $D_{1^{n-r}}\{\lambda\}$, which is a non-vanishing polynomial in $s_{\nu_1}, s_{\nu_2}, \ldots, s_{\nu_n}$.

Hence

$$a_r = \frac{D_{1^{n-r}}\{\lambda\}}{D_{1^n}\{\lambda\}}.$$

EXAMPLE. Take $n=4$ and the fundamental set as s_1, s_3, s_5, s_7. We have $\{\lambda\} = \{4321\}$ and $a_4 = \{4321\}/\{321\}$

$$= \frac{s_1^{10} - 15s_1^7 s_3 + 63s_1^5 s_5 - 225s_1^3 s_7 + 315s_1^2 s_3 s_5 - 175s_1 s_3^3 + 225s_3 s_7 - 189s_5^2}{105(s_1^6 - 5s_1^3 s_3 + 9s_1 s_5 - 5s_3^2)}.$$

The denominator $\{321\}$ is of course read off from the character tables; it could be verified that it is

$$D_{1^3}\{4321\} = \left(\frac{1}{4!}\frac{\partial^4}{\partial s_1^4} + \frac{\partial^2}{\partial s_1 \partial s_3}\right)\{4321\}.$$

Also

$$a_3 = \frac{D_1\{4321\}}{\{321\}} = \frac{\dfrac{\partial}{\partial s_1}\{4321\}}{\{321\}} = \frac{2s_1^9 - 21s_1^6 s_3 + 63s_1^4 s_5 - 135s_1^2 s_7 + 126s_1 s_3 s_5 - 35s_3^3}{21(s_1^6 - 5s_1^3 s_3 + 9s_1 s_5 - 5s_3^2)}$$

and

$$a_2 = \frac{D_{1^2}\{4321\}}{\{321\}} = \frac{\dfrac{1}{2}\dfrac{\partial^2}{\partial s_1^2}\{4321\}}{\{321\}} = \frac{3(s_1^8 - 7s_1^5 s_3 + 14s_1^3 s_5 - 15s_1 s_7 + 7s_3 s_5)}{7(s_1^6 - 5s_1^3 s_3 + 9s_1 s_5 - 5s_3^2)}.$$

Finally the theorem gives

$$a_1 = \frac{D_{1^3}\{4321\}}{\{321\}} = \frac{\left(\dfrac{1}{3!}\dfrac{\partial^3}{\partial s_1^3} + \dfrac{\partial}{\partial s_3}\right)\{4321\}}{\{321\}}$$

which of course reduces to s_1.

5. Pólya's Theorem

Incidental to his discussion of the precise expression of a_r in terms of $s_1, s_3, \ldots, s_{2n-1}$, PÓLYA [7] proved some interesting results on two sets of n indeterminates. Here I shall prove these results by means of SCHUR functions. As in the proof of KAKEYA's theorem, this method excels in that it gives complete information as to the actual rational functions involved. I have shown previously [11] that if $s_k = \sum_{i=1}^{n} x_i^k$ and $t_k = \sum_{i=1}^{n} y_i^k$, then

$$\varphi(s_1+t_1, s_2+t_2, \ldots, s_m+t_m) = [\{0\}' D_0 + \{1\}' D_1 + \{2\}' D_2 + \{1^2\}' D_{1^2} + \cdots]\, \varphi(s_1, s_2, \ldots, s_m),$$

where $\{\lambda\}'$ denotes an S-function of $y_1, y_2, \ldots, y_n$. If every t_k is replaced by $-t_k$ then $\{\lambda\}'$ changes into its conjugate $\{\tilde{\lambda}\}'$ and acquires a negative sign if (λ) is a partition of an odd number. Hence, if $u_k = s_k - t_k$, then

$$\text{(II)} \quad \begin{cases} \varphi(u_1, u_2, \ldots, u_m) = [1 - \{1\}' D_1 + \{1^2\}' D_2 + \{2\}' D_{1^2} - \\ \qquad - \{1^3\}' D_3 - \{21\}' D_{21} - \{3\}' D_{1^3} + \cdots]\, \varphi(s_1, s_2, \ldots, s_m). \end{cases}$$

In particular, the functions

$$v_1 = u_1, \quad v_2 = \tfrac{1}{2}(u_1^2 + u_2), \quad v_3 = \tfrac{1}{6}(u_1^3 + 3u_1u_2 + 2u_3),$$
$$v_4 = \tfrac{1}{24}(u_1^4 + 3u_2^2 + 8u_1u_3 + 6u_1^2u_2 + 6u_4), \ldots$$

used by PÓLYA, are obtained from $\{1\}, \{2\}, \{3\}, \ldots$, by merely replacing every s_i by u_i, and can be written as

$$\text{(III)} \quad v_r = \{r\} - \{1\}'\{r-1\} + \{1^2\}'\{r-2\} - \{1^3\}'\{r-3\} + \cdots + (-1)^r\{1^r\}'.$$

PÓLYA'S THEOREM. *If* $R = \prod_{i=1}^{n}\prod_{j=1}^{n}(x_i - y_j)$ *and* $b_r = \sum y_1 y_2 \ldots y_r$, *then*

$$\text{(i)} \quad V_n = \begin{vmatrix} v_1 & v_2 & \ldots & v_n \\ v_2 & v_3 & \ldots & v_{n+1} \\ \ldots & \ldots & \ldots & \ldots \\ v_n & v_{n+1} & \ldots & v_{2n-1} \end{vmatrix} = (-1)^{\frac{1}{2}n(n-1)} R;$$

(ii) R does not contain u_{2n};

(iii) *the* $2n+1$ *quantities* $R, Ra_1, Ra_2, \ldots, Ra_n; Rb_1, Rb_2, \ldots, Rb_n$ *are polynomials in* $u_1, u_2, \ldots, u_{2n}$.

PROOF. By elementary transformations of the determinant V_n in conjunction with relation (III) it is easily shown that

$$V_n = (-1)^{\frac{1}{2}n(n-1)} \begin{vmatrix} \{1\} - \{1\}' & \{2\} - \{2\}' & \ldots & \{n\} - \{n\}' \\ \{1^2\} - \{1^2\}' & \{21\} - \{21\}' & \ldots & \{n1\} - \{n1\}' \\ \ldots & \ldots & \ldots & \ldots \\ \{1^n\} - \{1^n\}' & \{21^{n-1}\} - \{21^{n-1}\}' & \ldots & \{n1^{n-1}\} - \{n1^{n-1}\}' \end{vmatrix}.$$

If $x_i = y_j$, then both these variables will vanish from every u_k and every v_r will be a function of $n-1$ x's and $n-1$ y's. By Lemma 1 the terms in the last row of the above determinant will be zero and so we can prove (i).

Next

$$\begin{vmatrix} \{1\} & \{2\} & \dots & \{n\} \\ \{2\} & \{3\} & \dots & \{n+1\} \\ \dots & \dots & \dots & \dots \\ \{n\} & \{n+1\} & \dots & \{2n-1\} \end{vmatrix} = (-1)^{\frac{1}{2}n(n-1)} \begin{vmatrix} \{n\} & \{n+1\} & \dots & \{2n-1\} \\ \{n-1\} & \{n\} & \dots & \{2n-2\} \\ \dots & \dots & \dots & \dots \\ \{1\} & \{2\} & \dots & \{n\} \end{vmatrix} = (-1)^{\frac{1}{2}n(n-1)} \{n^n\}$$

and so R is easily written down from the polynomial expression for $\{n^n\}$ by merely writing u_i for every s_i. Since the diagram of (n^n) has no removable $2n$-strip then, by Lemma 2, $\{n^n\}$ will not contain s_{2n} and hence R will not contain u_{2n}, which proves (ii).

From (II),

$$R = [1 - \{1\}' D_1 + \{1^2\}' D_2 + \{2\}' D_{1^2} - \dots + (-1)^n \{n^n\}' D_{n^n}] \{n^n\}$$

where $\{\lambda\}'$ is confined to those partitions with no part greater than n and with not more than n parts. Using Lemma 1, we have $\{1^r\}\{n^n\} = \{(n+1)^r n^{n-r}\}$. The determinant form of this S-function in terms of the a_r will be of order $n+1$, and its leading principal minor of order n will be $\{n^n\}$. The $(n+1)$th column will be zero except for the lowest member a_r. Since no (λ) has more than n parts the operator D_λ will affect only the first n columns and so

$$D_\lambda \{1^r\}\{n^n\} = D_\lambda \{(n+1)^r n^{n-r}\} = \{1^r\} D_\lambda \{n^n\}.$$

Hence

$$\{1^r\} R = [1 - \{1\}' D_1 + \{1^2\}' D_2 + \{2\}' D_{1^2} - \dots + (-1)^n \{n^n\}' D_{n^n}] \{(n+1)^r n^{n-r}\}.$$

If $\{(n+1)^r n^{n-r}\} = F(s_1, s_2, \dots, s_{2n})$, then the polynomial F can be found from group character tables or known combinatorial rules for the coefficients [12] and we have

$$a_r R = F(u_1, u_2, \dots, u_{2n}).$$

A trivial modification, merely changing u_i to $-u_i$, suffices to give the polynomial $b_r R$ and so (iii) is proved.

In conclusion, I remark that other formulae given by PÓLYA and PERRON can be written in convenient S-function forms. Thus $w_1 = 2s_1$, $w_2 = 2s_1^2$, $w_3 = \frac{1}{3}(4s_1^3 + 2s_3)$, ... are all given by $w_r = 2(\{r\} + \{r-1, 1\} + \{r-2, 1^2\} + \dots + \{1^r\})$ and

$$\begin{vmatrix} w_1 & 2 \\ w_3 & w_2 \end{vmatrix}, \quad \begin{vmatrix} w_1 & 2 & . \\ w_3 & w_2 & w_1 \\ w_5 & w_4 & w_3 \end{vmatrix}, \quad \begin{vmatrix} w_1 & 2 & . & . \\ w_3 & w_2 & w_1 & 2 \\ w_5 & w_4 & w_3 & w_2 \\ w_7 & w_6 & w_5 & w_4 \end{vmatrix}, \dots$$

are numerical multiples of $\{21\}$, $\{321\}$, $\{4321\}$, ... Also

$$\begin{vmatrix} w_1 & w_2 & \dots & w_n \\ w_2 & w_3 & \dots & w_{n+1} \\ \dots & \dots & \dots & \dots \\ w_n & w_{n+1} & \dots & w_{2n-1} \end{vmatrix}$$

for successive values of n, becomes numerical multiples of $\{1\}$, $\{1\}\{21\}$, $\{21\}\{321\}$, $\{321\}\{4321\}$, ...

References

[1] KAKEYA, S.: On fundamental systems of symmetric functions. I, II. Jap. J. Math. **2**, 69—80 (1925); **4**, 77—85 (1927). — [2] LAGUERRE, E.: Sur un problème d'algèbre. Bull. Soc. Math. France **5**, 26—30 (1877); Oeuvres I, S. 119—122. — [3] BORCHARDT, C. W.: Über eine Eigenschaft der Potenzsummen ungerader Ordnung. Monatsber. Akad. Berlin **1857**, 301—311; Ges. Werke, S. 107—118. — [4] VAHLEN, K. T.: Über Fundamentalsysteme für symmetrische Funktionen. Acta math. **23**, 91—120 (1900). — [5] LUDWIG, B. v.: Ein neues Fundamentalsystem für symmetrische Funktionen. Math. Ann. **83**, 67—69 (1921). — [6] NAKAMURA, K.: On the representation of symmetric functions by power sums which form the fundamental system. Jap. J. Math. **4**, 87—92 (1927). — [7] PÓLYA, G.: Remarques sur un problème d'algèbre étudié par Laguerre. J. Math. pur. appl. **31**, 37—47 (1951). — [8] PERRON, O.: Über die Abhängigkeit von Potenzsummen und einen Satz von PÓLYA. Math. Z. **63**, 19—30 (1955). — [9] PERRON, O.: Über Potenzsummen. Math. Z. **64**, 103—114 (1956). — [10] LITTLEWOOD, D. E.: The theory of group characters, 2nd Ed. Oxford 1950. — [11] FOULKES, H. O.: Differential operators associated with S-functions. J. London Math. Soc. **24**, 136—143 (1949). — [12] BOERNER, M.: Darstellungen von Gruppen, p. 173—176. Berlin-Göttingen-Heidelberg 1955.

University College, Swansea, Wales (Great Britain)

(Eingegangen am 3. Februar 1956)

ZADDACH, A.
Math. Zeitschr. Bd. 65, S. 353—388 (1956)

Über Anti-Fano-Ebenen

Von
ARNO ZADDACH

Einleitung

Als *Anti-*FANO*-Ebenen* sollen hier solche projektiven Ebenen bezeichnet werden, in denen jedes Viereck kollineare Diagonalpunkte besitzt, also gegen das Axiom von FANO verstößt. In ihnen gilt trivialerweise der Satz vom vollständigen Vierseit.

Wird umgekehrt für eine Ebene der Satz vom vollständigen Vierseit vorausgesetzt, so ist sie nach [9] entweder eine Anti-FANO-Ebene oder aber eine Ebene, in der das FANO-Axiom erfüllt ist. Für den letzteren Fall hat MOUFANG die Gültigkeit des kleinen DESARGUESschen Satzes bewiesen und aus diesem eine Reihe von Folgerungen gezogen:

(α) *Jede von einem nichtausgearteten Punktquadrupel erzeugte Unterebene* (in [4] als 8-parametriges Netz bezeichnet) *ist eine Ebene über einem Primkörper;*

(β) *jede Unterebene, die von dem eben genannten Quadrupel und einem weiteren Punkt auf einer Verbindungsgerade erzeugt wird* (in [5]: 9-parametriges Netz), *ist eine Ebene über einer einfachen Erweiterung eines Primkörpers;*

(γ) *jede Unterebene, die von dem genannten Quadrupel und zwei weiteren Punkten einer Verbindungsgerade erzeugt wird* (in [6], [7]: 10-parametriges Netz), *ist eine Ebene über einem Schiefkörper;*

(δ) *die ganze Ebene ist eine Ebene über einem Alternativkörper* [8].

Im Fall der Anti-FANO-Ebene ist natürlich (α) richtig: Das nichtausgeartete Quadrupel erzeugt die bekannte Sieben-Punkte-Ebene über dem Primkörper der Charakteristik 2. Zweck der vorliegenden Arbeit ist es, auch (β) zu beweisen. Die Frage, ob in jeder Anti-FANO-Ebene der kleine DESARGUESsche Satz und damit auch (γ) und (δ) gilt, bleibt weiterhin unentschieden.

Für die Anregung zu dieser Arbeit sowie für wertvolle Hinweise bin ich Herrn G. PICKERT zu großem Dank verpflichtet.

§1. Ternärkörper

Dem Begriff der *projektiven Ebene* werden die Inzidenzaxiome und ein Reichhaltigkeitsaxiom zugrunde gelegt:

Je zwei verschiedene Punkte A und B lassen sich durch genau eine Gerade AB verbinden;

je zwei verschiedene Geraden AB und CD schneiden sich in genau einem Punkt $AB \cap CD$;

es gibt vier verschiedene Punkte, von denen keine drei kollinear sind.

Eine *affine Ebene* erhält man hieraus, indem man eine Gerade und die auf ihr liegenden Punkte als *uneigentlich* bezeichnet. Ausdrücke wie „*parallel*" sind dann im üblichen Sinn zu verstehen. Den eigentlichen Teil der affinen Ebene kann man nach [3] durch Einführung eines Koordinatensystems folgendermaßen algebraisieren (die hier benutzte Formulierung stammt aus [10], S. 35 ff.).

Zwei Geraden mit eigentlichem Schnittpunkt O werden zur x- und y-*Achse* erklärt; ihre uneigentlichen Punkte seien U_0 bzw. U_∞. Außerdem wird noch ein eigentlicher Punkt E, der nicht auf den Achsen liegt, ausgezeichnet. Jeder eigentliche Punkt P der Ebene kann dann eineindeutig festgelegt werden durch sein *Koordinatenpaar* (x_P, y_P), das bestimmt wird durch

$$x_P = (P\,U_\infty \cap O\,E)\,U_0 \cap O\,U_\infty$$
$$y_P = P U_0 \cap O\,U_\infty .$$

Die Koordinaten sind also eigentliche Punkte der y-Achse. Ihre Menge soll $\mathfrak{K}$ genannt und nun zu einer algebraischen Struktur gemacht werden. Für die Punkte O und $O\,U_\infty \cap E\,U_0$ wird dabei 0 und 1 geschrieben, die übrigen Elemente aus $\mathfrak{K}$ werden mit kleinen griechischen oder lateinischen Buchstaben bezeichnet.

Um jede eigentliche Gerade durch eine Gleichung für die Koordinaten ihrer Punkte darstellen zu können, definiert man in $\mathfrak{K}$ eine *ternäre Verknüpfung* T mittels der geometrischen Konstruktionsvorschrift

$$T(\varkappa, \xi, \alpha) = \{[(\varkappa\,U_0 \cap E\,U_\infty)\,O \cap U_0\,U_\infty]\,\alpha \cap (\xi\,U_0 \cap O\,E)\,U_\infty\}\,U_0 \cap O\,U_\infty .$$

Dann ist die gewünschte Darstellung der eigentlichen Geraden entweder von der Form $y = T(\varkappa, x, \alpha)$ oder, falls es sich um eine Parallele zur y-Achse handelt, von der Form $x = \xi$. Alle Geraden der erstgenannten Art mit festem $\varkappa$ bilden ein Parallelenbüschel, dessen uneigentlicher Punkt mit $U_\varkappa$ bezeichnet wird.

Aus der Definition der Verknüpfung T ergeben sich sofort vier einfache Rechenregeln:

$$T(\varkappa, 1, 0) = \varkappa$$
$$T(1, \xi, 0) = \xi$$
$$T(\varkappa, 0, \alpha) = \alpha = T(0, \xi, \alpha) .$$

Außerdem gilt wegen der Inzidenzaxiome:

Zu vorgegebenem $\varkappa_1, \varkappa_2, \alpha_1, \alpha_2$ *mit* $\varkappa_1 \neq \varkappa_2$ *gibt es genau ein* ξ *mit*

$$T(\varkappa_1, \xi, \alpha_1) = T(\varkappa_2, \xi, \alpha_2);$$

zu vorgegebenem $\xi_1, \xi_2, \eta_1, \eta_2$ *mit* $\xi_1 \neq \xi_2$ *gibt es genau ein Paar* $\varkappa, \alpha$ *mit*

$$\begin{cases} \eta_1 = T(\varkappa, \xi_1, \alpha) \\ \eta_2 = T(\varkappa, \xi_2, \alpha); \end{cases}$$

zu vorgegebenem $\eta, \varkappa, \xi$ *gibt es genau ein* α *mit*

$$\eta = T(\varkappa, \xi, \alpha) .$$

$\mathfrak{R}$ wird jetzt als der auf O, U_0, U_∞, E bezogene *Ternärkörper* der Ebene bezeichnet. Isomorphie zwischen den Ternärkörpern einer Ebene braucht im allgemeinen nicht zu bestehen.

Man erklärt weiter zwei binäre Verknüpfungen, Addition und Multiplikation genannt, als Sonderfälle der ternären Verknüpfung T.

Definition der Addition: $\xi + \alpha = T(1, \xi, \alpha)$. Die Elemente aus $\mathfrak{R}$ bilden bezüglich der Addition eine Loop mit dem neutralen Element 0.

Definition der Multiplikation: $\varkappa\xi = T(\varkappa, \xi, 0)$. Die von 0 verschiedenen Elemente aus $\mathfrak{R}$ bilden bezüglich dieser Multiplikation eine Loop mit dem neutralen Element 1, und alle Elemente erfüllen die Gleichung $0\varkappa = \varkappa 0 = 0$.

Ein mit T gebildeter Term soll *zerlegbar* genannt werden, wenn sich die Verknüpfung der Argumente auch mittels Addition und Multiplikation ausdrücken läßt. Insbesondere soll unter *Hauptzerlegbarkeit* eine Aussage der Form

$$T(\varkappa, \xi, \alpha) = \varkappa\xi + \alpha$$

und unter *Nebenzerlegbarkeit* eine Aussage von einer der Formen

$$T(\varkappa, \xi, \varkappa\alpha) = \varkappa(\xi + \alpha) \quad \text{oder} \quad T(\varkappa, \xi, \alpha\xi) = (\varkappa + \alpha)\xi$$

verstanden werden.

§ 2. Potenzsummen

Von jetzt ab wird von der betrachteten Ebene vorausgesetzt, daß sie eine *Anti-FANO-Ebene* ist. Diese Eigenschaft, auf die im folgenden mit der Abkürzung (AF) verwiesen wird, besagt (Fig. 1):

Wenn für sieben verschiedene Geraden $I, \ldots, VII$ *und sieben verschiedene Punkte* $P_1, \ldots, P_7$ *die Inzidenzen*

I	$P_1\ P_2\ P_3$
II	$P_1\ P_4\ P_5$
III	$P_1\ P_6\ P_7$
IV	$P_2\ P_4\ P_7$
V	$P_2\ P_5\ P_6$
VI	$P_3\ P_4\ P_6$
VII	$P_3\ P_5$

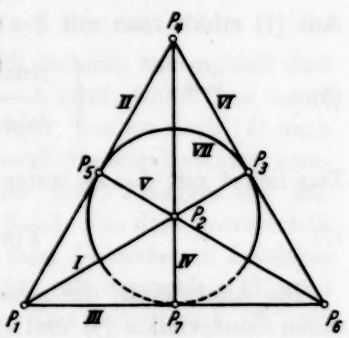


Fig. 1

gelten, dann besteht auch die Inzidenz $P_7 \in VII$. (Für das nichtausgeartete Viereck mit den Ecken P_1, P_2, P_4, P_6 bedeutet das die Kollinearität seiner Diagonalpunkte P_3, P_5, P_7.)

Aus Vierecken, die eine ausgezeichnete Lage bezüglich des Koordinatensystems besitzen, kann man leicht auch ohne Benutzung dieses allgemeinen Schemas Rechenregeln für *Anti-FANO-Ternärkörper* gewinnen. So äußert sich bei einem Parallelogramm die Eigenschaft (AF) darin, daß die Diagonalen

auch parallel sind, und das läßt sich bei einem solchen Parallelogramm, dessen eines Seitenpaar aus den Geraden $x=\xi$ und $x=\eta$ besteht, sofort algebraisch ausdrücken:

$$(1)\qquad \begin{array}{ll} & \left\{\begin{array}{l} T(\varkappa,\xi,\alpha)=T(\lambda,\xi,\gamma) \\ T(\varkappa,\xi,\beta)=T(\lambda,\xi,\delta) \\ T(\varkappa,\eta,\alpha)=T(\lambda,\eta,\delta) \end{array}\right. \\ \textit{folgt} & T(\varkappa,\eta,\beta)=T(\lambda,\eta,\gamma). \end{array}$$

(Damit ein nichtausgeartetes Parallelogramm vorlag, war für die Herleitung dieser Rechenregel die Voraussetzung $\xi\neq\eta$, $\varkappa\neq\lambda$ zu machen, die sich nachträglich als überflüssig herausstellt.) Durch Spezialisierung ergeben sich übersichtlichere Regeln. Man erhält zunächst mit $\lambda=0$:

$$(2)\qquad \left\{\begin{array}{ll} \textit{Aus} & T(\varkappa,\eta,\alpha)=T(\varkappa,\xi,\beta) \\ \hline \textit{folgt} & T(\varkappa,\eta,\beta)=T(\varkappa,\xi,\alpha). \end{array}\right.$$

Daraus weiter mit $\xi=0$:

$$(3)\qquad \textit{Aus}\quad T(\varkappa,\eta,\alpha)=\beta \quad \textit{folgt} \quad T(\varkappa,\eta,\beta)=\alpha.$$

Also insbesondere mit $\alpha=0$:

$$(4)\qquad T(\varkappa,\eta,\varkappa\eta)=0.$$

Unter Berücksichtigung von (4) bekommt man aus (2) mit $\beta=\varkappa\xi$:

$$(5)\qquad T(\varkappa,\eta,\varkappa\xi)=T(\varkappa,\xi,\varkappa\eta).$$

Aus (1) erhält man mit $\xi=0$:

$$(6)\qquad \left\{\begin{array}{ll} \textit{Aus} & T(\varkappa,\eta,\alpha)=T(\lambda,\eta,\beta) \\ \hline \textit{folgt} & T(\varkappa,\eta,\beta)=T(\lambda,\eta,\alpha). \end{array}\right.$$

Das liefert mit $\beta=\lambda\eta$ unter Berücksichtigung von (4):

$$(7)\qquad T(\varkappa,\eta,\lambda\eta)=T(\lambda,\eta,\varkappa\eta).$$

Die Ähnlichkeit, die zwischen den Formeln (2) und (6) sowie zwischen deren Sonderfällen (5) und (7) besteht, gibt Anlaß zu folgender Definition. Zwei Rechenregeln eines Anti-FANO-Ternärkörpers sollen *dual* zueinander heißen, wenn die eine aus der andern dadurch hervorgeht, daß in jedem T-Term das erste mit dem zweiten Argument vertauscht wird. Dabei soll eine Summe *nicht* als spezieller T-Term aufgefaßt werden, wohl aber ein Produkt gemäß der Definition der Multiplikation, so daß seine Faktoren auch zu vertauschen sind. (Im Anschluß an die Vertauschung ist natürlich eine Bezeichnungsänderung der Variablen zugelassen.) — *Dann lassen sich aus dualen Voraussetzungen duale Folgerungen ziehen, sofern während der Rechnung*

nur solche Eigenschaften von T benutzt werden, deren Dualisierungen ebenfalls Eigenschaften von T sind, also insbesondere von der Definitionsgleichung der Addition kein Gebrauch gemacht wird.

Zu der Regel (1) läßt sich das duale Gegenstück beweisen:

$$(8)\qquad \begin{array}{ll} \textit{Aus} & \left\{\begin{array}{l} T(x,\xi,\alpha)=T(x,\eta,\gamma) \\ T(x,\xi,\beta)=T(x,\eta,\delta) \\ T(\lambda,\xi,\alpha)=T(\lambda,\eta,\delta) \end{array}\right. \\ \textit{folgt} & T(\lambda,\xi,\beta)=T(\lambda,\eta,\gamma). \end{array}$$

Beweis. Mit geänderter Bezeichnung der dritten Argumente lautet (1):

$$\begin{array}{ll} \textit{Aus} & \left\{\begin{array}{l} T(x,\xi,\pi)=\beta=T(\lambda,\xi,\sigma) \\ T(x,\xi,\varrho)=\alpha=T(\lambda,\xi,\tau) \\ T(x,\eta,\pi)=\delta=T(\lambda,\eta,\tau) \\ T(x,\eta,\varrho)=\gamma \end{array}\right. \\ \textit{folgt} & \gamma=T(\lambda,\eta,\sigma). \end{array}$$

Durch Anwendung von (3) auf alle T-Terme läßt sich dies in (8) umwandeln.

Satz 1. *Für alle Anti-*FANO*-Ternärkörper gilt:*
Jedes Element genügt der Gleichung $x+x=0$;
die Addition ist kommutativ und assoziativ;
jedes Element der multiplikativen Loop hat ein Inverses;
die Multiplikation ist potenzassoziativ.

Beweis. Die Gleichung $x+x=0$ geht aus (4) dadurch hervor, daß dort $x=1$ genommen wird. Wenn man auch in (5) $x=1$ setzt, erhält man sofort die Kommutativität der Addition aller Ternärkörper, woraus nach [*1*] auch die Assoziativität folgt. In der Sprache der Gewebetheorie bedeutet kommutative Addition, daß die „THOMSEN-*Figur für jedes 3-Gewebe mit kollinearen Trägerpunkten*" geschlossen ist (vgl. [*10*], S. 59). Die dazu geometrisch duale Aussage, daß die „*Sechseck-Bedingung für jedes 3-Gewebe mit beliebigen Trägerpunkten*" erfüllt ist (vgl. [*10*], S. 54 und 154), gilt demnach auch, weil die Eigenschaft (AF) geometrisch selbstdual ist. Das liefert nach [*2*] und [*10*], S. 244, gerade die Behauptungen über die Multiplikation. (Bei den beiden zuletzt genannten Schließungssätzen handelt es sich um die „*kleinen Sätze von* PAPPOS".)

Auf die Gültigkeit der Sechseck-Bedingung wird später noch mehrfach Bezug genommen. Sie lautet ausführlich (Fig. 2):

Es seien drei verschiedene (Träger-) Punkte Q_1, Q_2, Q_3 gegeben, ferner neun verschiedene Geraden I, ..., IX, von denen jede durch genau einen der Punkte Q_i geht, und zwar die ersten drei durch Q_1, die nächsten drei

durch Q_2 und die letzten drei durch Q_3. Wenn dann für diese Geraden und sieben Punkte $P_0, \ldots, P_6$ die Inzidenzen

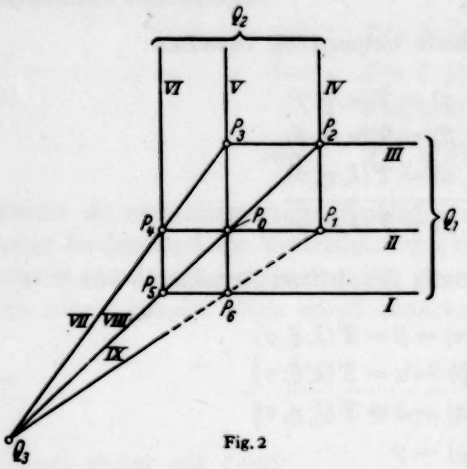


Fig. 2

I	$P_6\ P_5$
II	$P_1\ P_4\ P_0$
III	$P_2\ P_3$
IV	$P_2\ P_1$
V	$P_3\ P_6\ P_0$
VI	$P_4\ P_5$
VII	$P_4\ P_3$
VIII	$P_5\ P_2\ P_0$
IX	P_1

gelten, dann besteht auch die Inzidenz $P_6 \in IX$. (Diese Folgerung ist trivial im Entartungsfall $I = II = III$, $IV = V = VI$, $VII = VIII = IX$.)

In den Anwendungen werden stets Q_1 und Q_2 uneigentliche Punkte, Q_3 ein eigentlicher Punkt sein.

Nach Satz 1 kann man mit einem von 0 und 1 verschiedenen *erzeugenden Element* a aus $\mathfrak{R}$ und ganzen Zahlen e_i *Potenzsummen* bilden, d.h. Ausdrücke der Form

$$a^{e_1} + a^{e_2} + \cdots + a^{e_s}. \tag{9}$$

Bei dieser Schreibweise soll, falls die Summe überhaupt aus mehr als einem Glied besteht, stets $e_1 < e_2 < \cdots < e_s$ sein. Für eine solche a-Potenzsumme wird auch $\sum_\nu \delta_\nu a^\nu$ geschrieben, womit gemeint ist, daß der Index ν von $-\infty$ bis $+\infty$ läuft und endlich viele der Koeffizienten δ_ν gleich 1, die übrigen gleich 0 sind.

Läßt sich ein Element f aus $\mathfrak{R}$ als Potenzsumme (9) darstellen, so soll die maximale Exponentendifferenz deren *Länge* genannt werden:

$$Lg\, f = e_s - e_1. \tag{10}$$

Die Länge ist also eine nichtnegative ganze Zahl. Sie braucht durch das Ternärkörperelement f nicht eindeutig bestimmt zu sein, da möglicherweise die Darstellung als a-Potenzsumme in verschiedener Form erfolgen kann. Wo aber später mit (10) gearbeitet wird, ist immer klar erkennbar, auf welche Darstellung Bezug genommen wird. — Der Potenzsumme, die aus null Gliedern besteht, wird keine Länge zugeordnet, doch soll verabredet werden, daß diese leere Potenzsumme jede Bedingung $Lg\, f < n$ erfüllt.

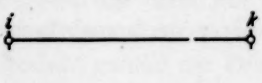


Fig. 3

Die verwendeten schematischen Darstellungen sind folgendermaßen zu verstehen:

In Fig. 3 bedeutet das obere Zeichen eine Potenzsumme $a^i+\cdots+a^k$ $(i\leq k)$, das untere Zeichen eine Potenzsumme $\delta a^i+\cdots+a^k$ ($\delta=0$ oder 1), d.h. die Potenz a^k tritt tatsächlich auf, und die Länge ist höchstens gleich $k-i$. Über etwaige Zwischenglieder wird in beiden Fällen nichts ausgesagt.

Die Buchstaben f, g, h, p, q werden im folgenden ausschließlich zur Bezeichnung von a-Potenzsummen verwendet.

§ 3. Multiplikation der Potenzsummen

Mit der Regel (1) und ihren Sonderfällen wird jetzt wiederholt die Zerlegbarkeit spezieller $\boldsymbol{T}$-Terme bewiesen. Dabei geschieht die Anwendung von (1) immer wie folgt: In den ersten drei Zeilen werden links und rechts Terme eingesetzt, für welche die Zerlegbarkeit schon bekannt ist (z.B. auf Grund einer Induktionsvoraussetzung), so daß die Richtigkeit dieser Gleichungen nachgeprüft werden kann. Ferner wird auch in der vierten Zeile rechts ein zerlegbarer Term eingesetzt. Damit gewinnt man dann die Zerlegbarkeit der linken Seite der letzten Zeile. — Entsprechend wird etwa die Regel (2) benutzt.

Als erstes Ergebnis wird nach diesem Verfahren die Hauptzerlegbarkeit

$$\boldsymbol{T}(a^\mu, a^{k-\mu}, h) = a^k + h \tag{11}$$

hergeleitet; dabei seien μ und k ganze Zahlen und $h=a^s+\cdots+a^t$.

Beweis. Es genügt die Betrachtung nichtnegativer Exponenten μ, da sich diese Form des $\boldsymbol{T}$-Terms notfalls durch Einführung eines neuen erzeugenden Elements $\bar{a}=a^{-1}$ und Übergang zu $\bar{a}$-Potenzen und $\bar{a}$-Potenzsummen herstellen läßt. — Die Behauptung ist nach der Definition der Addition richtig für $\mu=0$.

Erste Induktionsannahme. Die Behauptung sei schon als richtig erkannt für $\mu<m(>0)$. Sie soll nun für $\mu=m$ bewiesen werden. Man definiert als *Überhang von h über a^k* eine ganze Zahl $\varrho\geq 0$ durch

$$\varrho = \operatorname{Max}(t-s, k-s, t-k).$$

Bei $\varrho=0$ besteht h nur aus der einzigen Potenz a^k; dafür ist nach (4) die Behauptung richtig.

Zweite Induktionsannahme. Die Behauptung (11) sei bei $\mu=m$ schon bewiesen für $\varrho<r(>0)$. Nun soll (11) auch für $\varrho=r$ hergeleitet werden. Dabei werden drei Fälle unterschieden.

Fall A: $t<k$. Hier ist $r=k-s$. Es wird zunächst der Unterfall $r=1$ behandelt. h besteht dann nur aus der einzigen Potenz a^{k-1}. Man wendet (6) wie folgt an, bestätigt die erste Gleichung mit (4) und kann in der Schlußgleichung den rechten Term wegen der ersten Induktionsvoraussetzung zerlegen:

$$\boldsymbol{T}(a^m, a^{k-m}, a^k) = \boldsymbol{T}(a^{m-1}, a^{k-m}, a^{k-1}) = 0$$

$$\boldsymbol{T}(a^m, a^{k-m}, a^{k-1}) = \boldsymbol{T}(a^{m-1}, a^{k-m}, a^k) \quad = a^k + a^{k-1}.$$

Im Unterfall $r > 1$ ist (1) folgendermaßen anwendbar:

$$T(a^m, a^{s-m+1}, h + a^k + a^{k-1} + a^{s+1} + a^s) = T(a^{m-1}, a^{s-m+1}, h + a^k + a^{k-1})$$
$$T(a^m, a^{s-m+1}, h) \qquad\qquad = T(a^{m-1}, a^{s-m+1}, h + a^{s+1} + a^s)$$
$$T(a^m, a^{k-m}, h + a^k + a^{k-1} + a^{s+1} + a^s) \quad = T(a^{m-1}, a^{k-m}, h + a^{s+1} + a^s)$$

$$T(a^m, a^{k-m}, h) = T(a^{m-1}, a^{k-m}, h + a^k + a^{k-1}) = a^k + h.$$

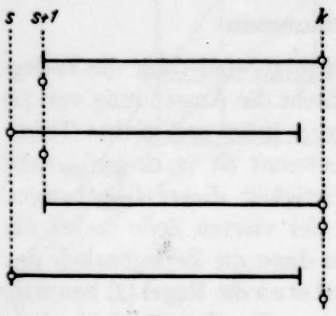


Fig. 4

Rechts fallen alle Terme unter die erste Induktionsvoraussetzung. Links werden in den ersten drei Zeilen die Terme durch die zweite Induktionsvoraussetzung erfaßt, wie man leicht aus dem Schema der Fig. 4 erkennt; in jeder Zeile steht dort das Produkt aus dem ersten und zweiten Argument des jeweiligen Terms unter dem dritten Argument.

Fall B: $t = k$. Hier ist h so beschaffen, daß $h + a^k$ zum Fall A gehört. Man hat deshalb $T(a^m, a^{k-m}, h + a^k) = h$ und kann nun gemäß (3) h mit $h + a^k$ vertauschen.

Fall C: $t > k$. Hier kommt man durch Einsetzen passender Größen in (2) zum Ziele:

$$T(a^m, a^{k-m}, h + a^t + a^k) = T(a^m, a^{t-m}, h)$$
$$T(a^m, a^{k-m}, h) \qquad = T(a^m, a^{t-m}, h + a^t + a^k) = a^k + h.$$

Für den rechten Term der oberen Gleichung (erste Zeile des Schemas Fig. 5) ist der Überhang von h über a^t maßgebend; bei $s \leq k$ ist dieser gleich r (der Term fällt dann unter Fall B), bei $s > k$ ist er kleiner als r. — Für den rechten Term der unteren Gleichung (zweite Zeile des Schemas) ist der Überhang von $h + a^t + a^k$ über a^t bei $s < k$ und bei $s > k$ gleich r (der Term fällt dann unter Fall A), bei $s = k$ kleiner als r. — Für den linken Term der oberen Gleichung (dritte Zeile des Schemas) ist der Überhang von $h + a^t + a^k$ über a^k offenbar kleiner als der Überhang von h über a^k, also kleiner als r.

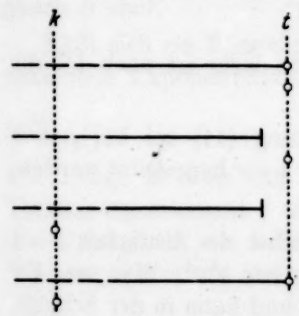


Fig. 5

Damit ist (11) vollständig bewiesen. Verallgemeinerungen dieser Formel hängen bereits mit Distributivitäts-Aussagen zusammen.

$$(12) \qquad a^\mu g \equiv a^\mu \sum_\nu \delta_\nu a^{\nu-\mu} = \sum_\nu \delta_\nu a^\nu,$$

$$(13) \qquad T(a^\mu, g, h) = a^\mu g + h.$$

Beweis. Es genügt wieder die Betrachtung nichtnegativer Exponenten μ. Die beiden Behauptungen sind richtig für $Lg\, g = 0$, denn (13) hat dann die Form (11).

Erste Induktionsannahme. (12) und (13) seien schon bewiesen für alle g mit $Lg\, g < n$ (>0). Es soll nun zunächst das Distributivgesetz (12) auch für $Lg\, g = n$ hergeleitet werden. Bei $\mu = 0$ ist nichts zu beweisen.

Zweite Induktionsannahme. (12) sei bei $Lg\, g = n$ bereits als richtig nachgewiesen für $0 \leq \mu < m$. Es soll nun gezeigt werden, daß dann auch das Produkt $a^m g$ distributiv ausmultiplizierbar ist.

Der Anfang der Potenzsumme g (die mindestens aus zwei Gliedern besteht) möge sein

$$g = a^\sigma + a^\tau + \cdots.$$

Dann wird gesetzt

$$g_1 = \sum_{\nu=\sigma}^{\tau-1} a^\nu.$$

Diese Potenzsumme g_1 hat offenbar eine kleinere Länge als g; man kann daher auf Grund der ersten Induktionsvoraussetzung das Produkt $a g_1$ distributiv ausmultiplizieren und erhält

$$a g_1 = \sum_{\nu=\sigma+1}^{\tau} a^\nu,$$

so daß

$$g_1 + a g_1 = a^\sigma + a^\tau$$

ist. Der noch nicht berücksichtigte Teil von g, der möglicherweise aus null Gliedern besteht, hat eine kleinere Länge als g und kann deshalb nach der ersten Induktionsvoraussetzung als $a g_2$ geschrieben werden. Zu

$$g = g_1 + a g_1 + a g_2 \tag{14}$$

wird nun

$$g^* = g_2 + a g_1 + a g_2 \tag{15}$$

gebildet; dabei gilt, wie man leicht feststellt, $Lg\, g^* < n$.

Man betrachtet jetzt das Parallelogramm, das durch das folgende Inzidenzschema für sieben Geraden und sieben Punkte beschrieben wird.

I	$U_0\ U_\infty$	U_0	$U_{a^{m-1}}$	U_{a^m}
II	$y = 0$	U_0	$(g_1, 0)$	$(g^*, 0)$
III	$y = a^m g_1 + a^m g_2$	U_0	$(g_2, a^m g_1 + a^m g_2)$	$(g, a^m g_1 + a^m g_2)$
IV	$y = \boldsymbol{T}(a^{m-1}, x, a^{m-1} g_1)$	$U_{a^{m-1}}$	$(g_1, 0)$	$(g, a^m g_1 + a^m g_2)$
V	$y = \boldsymbol{T}(a^{m-1}, x, a^{m-1} g^*)$	$U_{a^{m-1}}$	$(g^*, 0)$	$(g_2, a^m g_1 + a^m g_2)$
VI	$y = \boldsymbol{T}(a^m, x, a^m g_1)$	U_{a^m}	$(g_1, 0)$	$(g_2, a^m g_1 + a^m g_2)$
VII	$y = \boldsymbol{T}(a^m, x, a^m g^*)$	U_{a^m}	$(g^*, 0)$	

Von den hierin behaupteten Inzidenzen sind lediglich diejenigen nichttrivial, die in der letzten Spalte ab Zeile *IV* stehen. Man hat in Zeile *IV* den Term $\boldsymbol{T}(a^{m-1}, g, a^{m-1} g_1)$ auszuwerten. Zunächst formt man mit (5) um in $\boldsymbol{T}(a^{m-1}, g_1, a^{m-1} g)$, kann nun nach der ersten Induktionsvoraussetzung zerlegen in $a^{m-1} g_1 + a^{m-1} g$ und erhält daraus nach (14) und der zweiten

Induktionsvoraussetzung die behauptete y-Koordinate $a^m g_1 + a^m g_2$. — Bei Zeile V kann man sofort zerlegen und ähnlich wie eben (15) verwenden. — Bei Zeile VI ist ebenfalls der T-Term zerlegbar und damit die Inzidenz klar.

Man hat noch zu untersuchen, ob die Voraussetzungen für die Anwendbarkeit von (AF) erfüllt sind. Da die drei uneigentlichen Punkte wegen $a \neq 0, 1$ offenbar voneinander verschieden sind, ist das Parallelogramm genau dann ausgeartet, wenn die Geraden II und III zusammenfallen, also $a^m g_1 = a^m g_2$ gilt. Dann ist $a g_1 + a g_2 = 0$, und man hat nach (14)

$$a^m g = a^m (g_1 + 0) = a^m g_1 + a^m 0 = a^m g_1 + a^m (a g_1 + a g_2).$$

Hier ist aber rechts nach der ersten Induktionsvoraussetzung alles distributiv ausmultiplizierbar, denn gemäß der Definition von g_1 und g_2 gilt $Lg\,(a g_1 + a g_2) < Lg\, g$. — Man braucht daher das Parallelogramm nur im Nichtausartungsfall $a^m g_1 \neq a^m g_2$ heranzuziehen. (AF) liefert dann

$$T(a^m, g, a^m g^*) = a^m g_1 + a^m g_2.$$

Der T-Term ist nach Anwendung von (5) zerlegbar. Man erhält schließlich

$$a^m g = a^m g^* + a^m g_1 + a^m g_2.$$

Hier kann man rechts auf Grund der ersten Induktionsvoraussetzung alles distributiv ausmultiplizieren, also auch das Produkt $a^m g$.

Man muß jetzt noch (13) für $Lg\, g = n$ beweisen. Es sei

$$g = a^{i-m} + \cdots + a^{k-m}$$
$$h = a^s + \cdots + a^t.$$

Unter Verwendung dieser Exponenten wird der *Überhang von h über $a^m g$* definiert durch

$$\varrho = \mathrm{Max}\,(t - k + i - s, t - k, i - s, 0).$$

Zuerst wird die Lage bei $\varrho = 0$ betrachtet. Ist $t < k$, so gilt $Lg\, h < Lg\, g = n$, und da man nach dem eben Bewiesenen $h = a^m h'$ schreiben darf, kann man mittels (5) umwandeln und dann wegen $Lg\, h' = Lg\, h < n$ nach der ersten Induktionsvoraussetzung zerlegen:

$$T(a^m, g, h) = T(a^m, h', a^m g) = h + a^m g.$$

Ist aber $t = k$, so hat man nach dem soeben Festgestellten die Zerlegbarkeit von $T(a^m, g, h + a^m g)$, aus der man mit (3) die Zerlegbarkeit von $T(a^m, g, h)$ gewinnt.

Dritte Induktionsannahme. (13) sei bei $Lg\, g = n$ schon als richtig erkannt für $\varrho < r$ (> 0). Nun soll (13) auch für $\varrho = r$ bewiesen werden. Dazu wird wie beim Beweis von (11) eine Fallunterscheidung gemacht.

Fall A $(t < k)$ und Fall B $(t = k)$. Hier wird (2) folgendermaßen verwendet:

$$T(a^m, g, h + a^s + a^m g) = T(a^m, a^{s-m}, h)$$

$$T(a^m, g, h) = T(a^m, a^{s-m}, h + a^s + a^m g) = a^m g + h.$$

Die rechts stehenden Terme sind nach (11) zerlegbar. Das Schema der Fig. 6 bezieht sich auf die linken Gleichungsseiten. Der obere Term fällt unter die dritte Induktionsvoraussetzung: Bei Ersetzung von h durch $h+a^s+a^m g$ wird der Überhang über $a^m g$ am Potenzsummen-Anfang mindestens um 1 erniedrigt, während er von der Änderung am Potenzsummen-Ende unbeeinflußt bleibt.

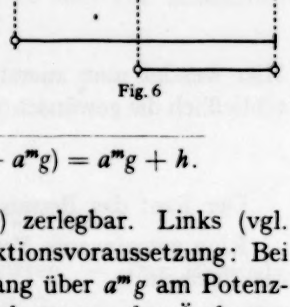


Fig. 6

Fall C ($t>k$) wird wie folgt mit (2) behandelt:

$$\begin{array}{ll} \boldsymbol{T}(a^m, g, h + a^i + a^m g) = \boldsymbol{T}(a^m, a^{i-m}, h) & \\ \hline \boldsymbol{T}(a^m, g, h) & = \boldsymbol{T}(a^m, a^{i-m}, h + a^i + a^m g) = a^m g + h. \end{array}$$

Die rechts stehenden Terme sind wieder nach (11) zerlegbar. Links (vgl. Fig. 7) fällt der obere Term unter die dritte Induktionsvoraussetzung: Bei Ersetzung von h durch $h+a^i+a^m g$ wird der Überhang über $a^m g$ am Potenzsummen-Ende mindestens um 1 erniedrigt, während er von der Änderung am Potenzsummen-Anfang unbeeinflußt bleibt, ganz gleich, ob der Exponent des Anfangsglieds a^s von h kleiner ist als der Exponent des Anfangsglieds a^i von $a^m g$ oder nicht.

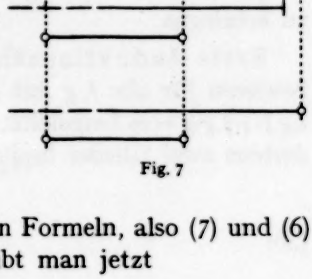


Fig. 7

Damit sind (12) und (13) vollständig bewiesen. Ganz entsprechend läßt sich auch zeigen

$$(16) \qquad f a^\mu = \Big(\sum_\nu \delta_\nu a^{\nu-\mu}\Big) a^\mu = \sum_\nu \delta_\nu a^\nu,$$

$$(17) \qquad \boldsymbol{T}(f, a^\mu, h) = f a^\mu + h.$$

Man benutzt zum Beweis einfach überall die dualen Formeln, also (7) und (6) an Stelle von (5) und (2). Statt (14), (15) schreibt man jetzt

$$f = f_1 + f_1 a + f_2 a$$
$$f^* = f_2 + f_1 a + f_2 a$$

mit $Lg f_1, Lg f_2, Lg f^* < Lg f$.

An die Stelle des Parallelogramms tritt das folgende Trapez (das man entsprechend dem damals Gesagten nur für $f_1 \neq f_2$, d.h. bei Nichtausartung benötigt).

I	$x = 0$	U_∞	$(0, 0)$	$(0, f_1 a^m + f_2 a^m)$
II	$x = a^{m-1}$	U_∞	$(a^{m-1}, f^* a^{m-1})$	$(a^{m-1}, f_1 a^{m-1})$
III	$x = a^m$	U_∞	$(a^m, f_1 a^m)$	$(a^m, f^* a^m)$
IV	$y = f^* x$	$(0, 0)$	$(a^{m-1}, f^* a^{m-1})$	$(a^m, f^* a^m)$
V	$y = f_1 x$	$(0, 0)$	$(a^{m-1}, f_1 a^{m-1})$	$(a^m, f_1 a^m)$
VI	$y = \boldsymbol{T}(f_2, x, f_1 a^m + f_2 a^m)$	$(0, f_1 a^m + f_2 a^m)$	$(a^{m-1}, f^* a^{m-1})$	$(a^m, f_1 a^m)$
VII	$y = \boldsymbol{T}(f, x, f_1 a^m + f_2 a^m)$	$(0, f_1 a^m + f_2 a^m)$	$(a^{m-1}, f_1 a^{m-1})$	

Die meisten Inzidenzen dieses Schemas sind leicht zu erkennen; die letzte in Zeile *VII* bestätigt man, indem man auf die Gleichung

$$\boldsymbol{T}(f_1, a^{m-1}, f a^{m-1}) = f_1 a^{m-1} + f a^{m-1},$$

die auf Grund der zu machenden Induktionsvoraussetzung gilt, zuerst (7) und dann (3) anwendet. — Nach (AF) folgt daher

$$\boldsymbol{T}(f, a^m, f_1 a^m + f_2 a^m) = f^* a^m.$$

Hier wendet man zuerst (3) und dann (7) an, darf nun zerlegen und erhält schließlich die gewünschte Aussage über die distributive Ausmultiplizierbarkeit:

$$f a^m = f^* a^m + f_1 a^m + f_2 a^m.$$

Der Rest des Beweises von (13) ist ohne Schwierigkeiten zu dualisieren.

Eine gemeinsame Verallgemeinerung von (12), (13) und (16), (17) bilden die Formeln

$$f g \equiv \Big(\sum_{\mu} \delta_\mu a^\mu\Big)\Big(\sum_{\nu} \varepsilon_\nu a^\nu\Big) = \sum_{\mu,\nu} \delta_\mu \varepsilon_\nu a^{\mu+\nu}, \tag{18}$$

$$\boldsymbol{T}(f, g, h) = f g + h. \tag{19}$$

Wegen des bereits Hergeleiteten ist der Beweis nur noch für $Lg\, f,\ Lg\, g > 0$ zu erbringen.

Erste Induktionsannahme. Die beiden Behauptungen seien schon bewiesen für alle f, g mit $Lg\, f + Lg\, g < m$ (>0). Nun soll zunächst (18) für $Lg\, f + Lg\, g = m$ hergeleitet werden. Da jede der beiden Potenzsummen mindestens zwei Glieder besitzt, kann man in der bekannten Weise setzen

$$\left\{\begin{aligned} f &= f_1 + f_1 a + f_2 a \\ g &= g_1 + a g_1 + a g_2 \\ g^* &= g_2 + a g_1 + a g_2 \end{aligned}\right. \tag{20}$$

mit $$Lg\, f_1,\ Lg\, f_2 < Lg\, f$$

und $$Lg\, g_1,\ Lg\, g_2,\ Lg\, g^* < Lg\, g.$$

Man betrachtet nun das folgende Parallelogramm

I	U_0 U_∞	U_{f_1}	U_{f_2}	U_f
II	$y = \boldsymbol{T}(f_1, x, f_1 g^*)$	U_{f_1}	$(g_1, f_1 g^* + f_1 g_1)$	$(g^*, 0)$
III	$y = \boldsymbol{T}(f_1, x, f_2 g + f_2 g_1 + f_1 g_2)$	U_{f_1}	$(g_2, f_2 g + f_2 g_1)$	$(g, f g_1 + f g_2)$
IV	$y = \boldsymbol{T}(f_2, x, f_1 g^* + f_2 g_1 + f_1 g_2)$	U_{f_2}	$(g_1, f_1 g^* + f_1 g_1)$	$(g, f g_1 + f g_2)$
V	$y = \boldsymbol{T}(f_2, x, f_2 g^*)$	U_{f_2}	$(g^*, 0)$	$(g_2, f_2 g + f_2 g_1)$
VI	$y = \boldsymbol{T}(f, x, f g_1 + f_1 g^* + f_1 g_1)$	U_f	$(g_1, f_1 g^* + f_1 g_1)$	$(g_2, f_2 g + f_2 g_1)$
VII	$y = \boldsymbol{T}(f, x, f g^*)$	U_f	$(g^*, 0)$	

Die Fälle, in denen die Voraussetzungen für die Anwendbarkeit von (AF) nicht erfüllt sind, werden vorher gesondert untersucht. Aus dem Zusammen-

fallen zweier uneigentlicher Punkte folgt mit (20) $f = f_1 = f_2$; dann ist

$$fg = (f_1 + 0)\,g = f_1 g + 0\,g = f_1 g + (f_1 a + f_2 a)\,g,$$

also distributiv ausmultiplizierbar wegen $Lg(f_1 a + f_2 a) < Lg\, f$. — Falls die uneigentlichen Punkte verschieden sind, also $f_1 \neq f_2$ gilt, so ist das Parallelogramm genau dann ausgeartet, wenn die Geraden II und III zusammenfallen. Dann ist

$$f_1(g^* + g_2) = f_2(g + g_1),$$

also

$$f_1(a g_1 + a g_2) = f_2(a g_1 + a g_2)$$

und deshalb $a g_1 + a g_2 = 0$. Daher kann man

$$fg = f(g_1 + 0) = fg_1 + f0 = fg_1 + f(a g_1 + a g_2)$$

wegen $Lg(a g_1 + a g_2) < Lg\, g$ distributiv ausmultiplizieren. — Das Parallelogramm wird demnach nur bei Nichtausartung benötigt.

Die Zerlegbarkeit ist überall durch die Induktionsvoraussetzung gesichert. Die Bestätigung der Inzidenzen erfordert lediglich in der rechten Spalte etwas Umrechnung. Man hat dafür unter Ausnutzung von (20) die Richtigkeit folgender Identitäten festzustellen.

Für Zeile III: $f_1 g + f_2 g + f_2 g_1 + f_1 g_2 = fg_1 + fg_2$.
Für Zeile IV: $f_2 g + f_1 g^* + f_2 g_1 + f_1 g_1 = fg_1 + fg_2$.
Für Zeile V: $f_2 g_2 + f_2 g^* = f_2 g + f_2 g_1$.
Für Zeile VI: $fg_2 + fg_1 + f_1 g^* + f_1 g_1 = f_2 g + f_2 g_1$.

Dann bekommt man gemäß (AF)

$$\boldsymbol{T}(f, g, fg^*) = fg_1 + fg_2.$$

Nach Anwendung von (5) auf den $\boldsymbol{T}$-Term darf man zerlegen und erhält schließlich die Gleichung

$$fg = fg^* + fg_1 + fg_2,$$

in der sich rechts nach der Induktionsvoraussetzung alle Produkte distributiv ausmultiplizieren lassen, was demnach auch für das Produkt fg zutrifft.

Damit ist (18) auch für $Lg\, f + Lg\, g = m$ als richtig erkannt. Von dieser Distributivität wird bei dem anschließenden Beweis von (19), der nun noch für $Lg\, f + Lg\, g = m$ zu erbringen ist, Gebrauch gemacht. — Es sei

$$f = a^i + \cdots + a^j$$
$$g = a^k + \cdots + a^l$$
$$h = a^s + \cdots + a^t.$$

Als Überhang von h über $a^j g$ ergibt sich

$$\varrho = \mathrm{Max}\big(t - (j + l) + (j + k) - s,\ t - (j + l),\ (j + k) - s,\ 0\big).$$

Bei $\varrho=0$ ist $\quad t-j-l\leqq 0$

und $\quad j+k-s\leqq 0,$

also $\quad Lg\,h=t-s\leqq l-k=Lg\,g.$

Man betrachtet nun, um die Behauptung für $\varrho=0$ zu beweisen, zuerst den Fall $t=j+l$. Dann gilt nach dem eben Bemerkten

$$Lg(g+ha^{-j})<Lg\,g.$$

Ferner ist offenbar

$$(21)\qquad \begin{cases} Lg(f+a^j)<Lg\,f \\ Lg(g+a^l)<Lg\,g. \end{cases}$$

Daher kann man (8) auf Grund der ersten Induktionsvoraussetzung folgendermaßen anwenden:

$$\begin{aligned} \boldsymbol{T}(f+a^j,g,0)&=\boldsymbol{T}(f+a^j,g+ha^{-j},h+fha^{-j})\\ \boldsymbol{T}(f+a^j,g,h)&=\boldsymbol{T}(f+a^j,g+ha^{-j},fha^{-j})\\ \boldsymbol{T}(f,g,0)&=\boldsymbol{T}(f,g+ha^{-j},fha^{-j})\\ \hline \boldsymbol{T}(f,g,h)&=\boldsymbol{T}(f,g+ha^{-j},h+fha^{-j})=fg+h. \end{aligned}$$

Für den Fall $t<j+l$ zieht man die Potenzsumme $h+a^{j+l}$ heran. Sie ordnet sich dem eben Erledigten unter und gewährleistet bei der folgenden Anwendung von (8) die Zerlegbarkeit des linken Terms der dritten Zeile, während man an den andern erforderlichen Stellen wegen (21) auf Grund der Induktionsvoraussetzung zerlegen kann.

$$\begin{aligned} \boldsymbol{T}(f+a^j,g,h+a^{j+l})&=\boldsymbol{T}(f+a^j,g+a^l,h+fa^l)\\ \boldsymbol{T}(f+a^j,g,h)&=\boldsymbol{T}(f+a^j,g+a^l,h+fa^l+a^{j+l})\\ \boldsymbol{T}(f,g,h+a^{j+l})&=\boldsymbol{T}(f,g+a^l,h+fa^l+a^{j+l})\\ \hline \boldsymbol{T}(f,g,h)&=\boldsymbol{T}(f,g+a^l,h+fa^l)=fg+h. \end{aligned}$$

Demnach ist (19) bei $Lg\,f+Lg\,g=m$ für $\varrho=0$ richtig.

Zweite Induktionsannahme. (19) sei bei $Lg\,f+Lg\,g=m$ bereits für $\varrho<r$ (>0) hergeleitet. Nun soll (19) auch für $\varrho=r$ bewiesen werden. Es wird wieder die übliche Fallunterscheidung gemacht und (1) angewendet.

Fall A und B $(t\leqq j+l)$. In dem folgenden Gleichungssystem kann man den linken Term der dritten Zeile nach der zweiten Induktionsvoraussetzung zerlegen, weil bei Ersetzung von h durch $h+a^jg+a^s$ der Überhang über a^jg am Potenzsummen-Anfang mindestens um 1 erniedrigt wird, während er von der Änderung am Potenzsummen-Ende nicht beeinflußt wird. Die sonst benötigten Zerlegbarkeiten sind wegen (21) durch die erste Induktionsvoraussetzung oder schon durch (17) gesichert.

$$\begin{aligned} \boldsymbol{T}(f,a^{s-j},h+a^jg+a^s)&=\boldsymbol{T}(f+a^j,a^{s-j},h+a^jg)\\ \boldsymbol{T}(f,a^{s-j},h)&=\boldsymbol{T}(f+a^j,a^{s-j},h+a^s)\\ \boldsymbol{T}(f,g,h+a^jg+a^s)&=\boldsymbol{T}(f+a^j,g,h+a^s)\\ \hline \boldsymbol{T}(f,g,h)&=T(f+a^j,g,h+a^jg)=fg+h. \end{aligned}$$

Der Fall C ($t > j + l$) ist zu erledigen mit

$$\begin{aligned} \boldsymbol{T}(f, a^{t-j}, h + a^j g + a^t) &= \boldsymbol{T}(f + a^j, a^{t-j}, h + a^j g) \\ \boldsymbol{T}(f, a^{t-j}, h) &= \boldsymbol{T}(f + a^j, a^{t-j}, h + a^t) \\ \boldsymbol{T}(f, g, h + a^j g + a^t) &= \boldsymbol{T}(f + a^j, g, h + a^t) \\ \hline \boldsymbol{T}(f, g, h) &= \boldsymbol{T}(f + a^j, g, h + a^j g) = fg + h. \end{aligned}$$

Der linke Term der dritten Zeile fällt unter die zweite Induktionsvoraussetzung, weil bei Ersetzung von h durch $h + a^j g + a^t$ der Überhang über $a^j g$ am Potenzsummen-Ende mindestens um 1 erniedrigt wird, während er von der Änderung am Potenzsummen-Anfang unbeeinflußt bleibt, gleichgültig, ob der Exponent s kleiner ist als der Exponent $j + k$ oder nicht. Die übrigen sechs benötigten Zerlegbarkeiten liefert die erste Induktionsvoraussetzung oder schon (17).

Damit ist der Induktionsbeweis von (18) und (19) beendet. Mit (18) ist aber gezeigt, daß der Bereich aller Elemente aus $\mathfrak{R}$, die durch a-Potenzsummen darstellbar sind, bezüglich der Addition und Multiplikation von $\mathfrak{R}$ einen kommutativen Ring bildet. Da es in $\mathfrak{R}$ keine Nullteiler gibt, hat man so den

Satz 2. *In einem Anti-*FANO*-Ternärkörper ist der Bereich aller Potenzsummen eines Elementes $a \neq 0$ ein Integritätsbereich.*

Endliche Integritätsbereiche sind Körper. Die in der Einleitung ausgesprochene Behauptung (β) ist daher bereits für den Fall bewiesen, daß die dort erwähnte Unterebene endlich ist. Algebraisch bedeutet das:

Satz 3. *Jeder endliche Anti-*FANO*-Ternärkörper, der von einem Element erzeugt wird, ist ein Körper der Charakteristik 2.*

Im folgenden wird deshalb nur noch der Fall betrachtet, daß der Integritätsbereich der a-Potenzsummen unendlich ist. Dann gibt es für jedes Element aus $\mathfrak{R}$ höchstens eine Darstellung als a-Potenzsumme. Aus zwei formal verschiedenen Darstellungen eines Elements folgt nämlich eine Gleichung

$$a^i + \cdots + a^k = 0;$$

dabei hat diese Potenzsumme wegen $a \neq 0, 1$ mindestens zwei Glieder mit Exponenten $i \leq k - 2$. Durch Multiplikation mit passenden a-Potenzen erhält man daraus

$$a^{-1} = \delta_n a^0 + \cdots + \delta_1 a^{n-1} + a^n, \tag{22}$$

$$a^{n+1} = 1 + \cdots + \delta_1 a^n, \tag{23}$$

mit $\delta_\nu = 0$ oder 1 und $n \geq 1$. Aus (22) entnimmt man, daß sich jede a-Potenz mit negativem Exponenten auf Potenzen mit nichtnegativen Exponenten zurückführen läßt, und nach (23) kann man jede a-Potenz mit einem Exponenten $> n$ ausdrücken durch eine Potenzsumme mit Exponenten zwischen 0 und n. Von diesen gibt es aber nur endlich viele.

Von jetzt an ist wegen der Eindeutigkeit der Darstellung jede a-Potenzsumme f mit $Lg\, f \geq 0$ verschieden von dem Ternärkörper-Element 0, so daß f nach Satz 1 ein Inverses $1/f$ besitzt. Andererseits wird für die nun herzuleitenden Rechenregeln, in denen ein Ausdruck $1/f$ vorkommt, grundsätzlich $f \neq 0$ vorausgesetzt, ohne daß es dort ausdrücklich erwähnt wird.

§4. Kürzungsregeln

Mit den Methoden des § 3 lassen sich auch die Linkskürzungsregel

$$\frac{1}{f}(f a^{\mu}) = a^{\mu} \tag{24}$$

und die Nebenzerlegbarkeit

$$T\left(\frac{1}{f}, f a^{\mu} + h, a^{\mu}\right) = \frac{1}{f} h, \tag{25}$$

die man nach (3) auch in der gleichwertigen Form

$$T\left(\frac{1}{f}, f a^{\mu} + h, \frac{1}{f} h\right) = a^{\mu} \tag{25'}$$

schreiben kann, sowie deren Verallgemeinerungen

$$\frac{1}{f}(f g) = g, \tag{26}$$

$$T\left(\frac{1}{f}, f g + h, g\right) = \frac{1}{f} h, \tag{27}$$

$$T\left(\frac{1}{f}, f g + h, \frac{1}{f} h\right) = g \tag{27'}$$

beweisen. Man hat dazu nur die Ecken der bei der Herleitung von (16), (17), (18), (19) benutzten Vierecke der Abbildung $(\xi, \eta) \to (\eta, \xi)$ zu unterwerfen. Auf die Beweise wird an späterer Stelle noch Bezug genommen; sie sollen deshalb in aller Ausführlichkeit dargestellt werden, obwohl nichts wesentlich Neues hinzukommt.

An Stelle des bei (1) betrachteten Parallelogramms wird hier ein solches verwendet, dessen eines Seitenpaar aus den Geraden $y = \alpha$ und $y = \beta$ besteht; wenn dann außerdem keine Seite oder Diagonale parallel zur y-Achse ist [weil diese Parallelität ja nur die schon bekannte Regel (2) liefern würde], läßt sich (AF) ausdrücken durch folgendes Gleichungssystem:

$$\begin{array}{ll} \textit{Aus} & \left\{ \begin{array}{l} T(\varkappa, \xi, \varphi) = T(\varkappa, \eta, \psi) = T(\lambda, \xi, \varrho) = T(\lambda, \eta, \sigma) = \alpha \\ T(\varkappa, \zeta, \varphi) = T(\varkappa, \vartheta, \psi) = T(\lambda, \vartheta, \varrho) = \beta \end{array} \right. \\ \textit{folgt} & \overline{\qquad\qquad T(\lambda, \zeta, \sigma) = \beta .} \end{array}$$

(Damit ein nichtausgeartetes Parallelogramm vorlag, war für die Herleitung die Voraussetzung $0 \neq \varkappa \neq \lambda \neq 0$, $\alpha \neq \beta$ zu machen, die sich nachträglich als überflüssig herausstellt.) Wendet man auf alle Terme (3) an, so bekommt

diese Rechenregel die an (1) erinnernde Form:

$$(28)\qquad \begin{array}{ll} & \left\{\begin{array}{l} \boldsymbol{T}(\varkappa,\xi,\alpha)=\boldsymbol{T}(\varkappa,\zeta,\beta)\\ \boldsymbol{T}(\varkappa,\eta,\alpha)=\boldsymbol{T}(\varkappa,\vartheta,\beta)\\ \boldsymbol{T}(\lambda,\xi,\alpha)=\boldsymbol{T}(\lambda,\vartheta,\beta)\end{array}\right.\\ \textit{folgt} & \boldsymbol{T}(\lambda,\eta,\alpha)=\boldsymbol{T}(\lambda,\zeta,\beta)\,.\end{array}$$

Da man hier offenbar jede der vier Gleichungen als Folgerung aus den drei andern ansehen kann, erhält man mit $\zeta=0=\alpha$ unter Berücksichtigung von (3) die Regel:

$$(29)\qquad \textit{Von den Gleichungen}\quad \left\{\begin{array}{r} \boldsymbol{T}(\varkappa,\vartheta,\varkappa\eta)=\beta\\ \boldsymbol{T}(\lambda,\vartheta,\lambda\xi)=\beta\\ \varkappa\xi=\beta\\ \lambda\eta=\beta\end{array}\right.$$

folgt aus je dreien die vierte.

Der Beweis von (24) und (25) ist nur für $\mu\geqq 0$ erforderlich, da man notfalls zu dem erzeugenden Element $\bar{a}=a^{-1}$ übergehen kann. — Die Behauptungen sind richtig für $Lg\,f=0$.

Erste Induktionsannahme. (24) und (25) seien bereits als richtig erkannt für alle f mit $Lg\,f<n\ (>0)$. Es soll nun für $Lg\,f=n$ zunächst (24) hergeleitet werden. Für $\mu=0$ ist nichts zu beweisen.

Zweite Induktionsannahme. (24) sei bei $Lg\,f=n$ schon als richtig erkannt für $0\leqq\mu<m$. Es folgt nun der Beweis für $\mu=m$. Dazu setzt man gemäß (14), (15)

$$f=f_1+f_1a+f_2a,$$
$$f^*=f_2+f_1a+f_2a$$

mit $Lg\,f_1$, $Lg\,f_2$, $Lg\,f^*<Lg\,f$. — Es sind nun die beiden Fälle $f_2\neq 0$ und $f_2=0$ zu unterscheiden.

Fall $f_2\neq 0$. Das folgende Trapez ist nichtausgeartet: Erstens sind natürlich I, II, III voneinander verschieden, und zweitens ist $V\neq VI$, weil f_1 und f_2 ihrer Definition gemäß verschiedene a-Potenzen enthalten und die Darstellung durch a-Potenzsummen jetzt eindeutig ist.

I	$y=0$	U_0	$(f_1a^m,0)$	$(f_2a^m,0)$
II	$y=a^{m-1}$	U_0	$(f_1a^{m-1}+f_1a^m,a^{m-1})$	$(f_2a^{m-1}+f_2a^m,a^{m-1})$
III	$y=a^m$	U_0	$(0,a^m)$	$(f^*a^m+f_1a^m,a^m)$
IV	$y=\boldsymbol{T}\left(\frac{1}{f^*},x,\frac{1}{f^*}(f_1a^m)\right)$	$(f_1a^m,0)$	$(f_2a^{m-1}+f_2a^m,a^{m-1})$	$(f^*a^m+f_1a^m,a^m)$
V	$y=\boldsymbol{T}\left(\frac{1}{f_1},x,a^m\right)$	$(f_1a^m,0)$	$(f_1a^{m-1}+f_1a^m,a^{m-1})$	$(0,a^m)$
VI	$y=\boldsymbol{T}\left(\frac{1}{f_2},x,a^m\right)$	$(f_2a^m,0)$	$(f_2a^{m-1}+f_2a^m,a^{m-1})$	$(0,a^m)$
VII	$y=\boldsymbol{T}\left(\frac{1}{f},x,\frac{1}{f}(f_2a^m)\right)$	$(f_2a^m,0)$	$(f_1a^{m-1}+f_1a^m,a^{m-1})$	

Die Bestätigung der Inzidenz in Zeile *IV*, Mitte, erfolgt mit

$$f_2 a^{m-1} + f_2 a^m = f^* a^{m-1} + f_1 a^m.$$

Für Zeile *VII*, Mitte, berücksichtigt man zunächst die Identität

$$f_1 a^{m-1} + f_1 a^m = f a^{m-1} + f_2 a^m$$

und zieht dann (29) heran: Es ist nach Induktionsvoraussetzung

$$\boldsymbol{T}\left(\frac{1}{f_2 a}, f_2 a^m + f a^{m-1}, \frac{1}{f_2 a}(f a^{m-1})\right) = a^{m-1}$$

$$\frac{1}{f}(f a^{m-1}) = a^{m-1}$$

$$\frac{1}{f_2 a}(f_2 a^m) = a^{m-1}.$$

Daher folgt

$$\boldsymbol{T}\left(\frac{1}{f}, f a^{m-1} + f_2 a^m, \frac{1}{f}(f_2 a^m)\right) = a^{m-1},$$

also die behauptete Inzidenz.

(AF) liefert nun wegen $f^* a^m + f_1 a^m = f a^m + f_2 a^m$ die Inzidenz

$$\boldsymbol{T}\left(\frac{1}{f}, f a^m + f_2 a^m, \frac{1}{f}(f_2 a^m)\right) = a^m.$$

Da man außerdem auf Grund der Induktionsvoraussetzung die beiden Gleichungen

$$\boldsymbol{T}\left(\frac{1}{f_2}, f_2 a^m + f a^m, \frac{1}{f_2}(f a^m)\right) = a^m$$

$$\frac{1}{f_2}(f_2 a^m) = a^m$$

zur Verfügung hat, kann man (29) anwenden und bekommt

$$\frac{1}{f}(f a^m) = a^m.$$

Fall $f_2 = 0$. Daß das folgende Trapez nichtausgeartet ist, erkennt man sofort wegen $I \neq II \neq III \neq I$ und $VI \neq VII$.

I	$y = 0$	U_0	$(f_1 a^m, 0)$	$(0, 0)$
II	$y = a^{m-1}$	U_0	$(f a^{m-1}, a^{m-1})$	$(0, a^{m-1})$
III	$y = a^m$	U_0	$(0, a^m)$	$(f a^m, a^m)$
IV	$y = \boldsymbol{T}\left(\frac{1}{f_1 a}, x, a^{m-1}\right)$	$(f_1 a^m, 0)$	$(0, a^{m-1})$	$(f a^m, a^m)$
V	$y = \boldsymbol{T}\left(\frac{1}{f_1}, x, a^m\right)$	$(f_1 a^m, 0)$	$(f a^{m-1}, a^{m-1})$	$(0, a^m)$
VI	$x = 0$	$(0, 0)$	$(0, a^{m-1})$	$(0, a^m)$
VII	$y = \frac{1}{f} x$	$(0, 0)$	$(f a^{m-1}, a^{m-1})$	

Die behaupteten Inzidenzen sind mit der Induktionsvoraussetzung und der Identität $f = f_1 a + f_1$ leicht zu bestätigen. (AF) liefert deshalb auch hier

$$\frac{1}{f}(f a^m) = a^m.$$

Damit ist (24) für $Lg\, f = n$ hergeleitet. Bei dem jetzt folgenden Beweis von (25) kann $h \neq 0$ angenommen werden, da sonst (25') mit (24) zusammenfällt. Es sei

$$f = a^{i-m} + \cdots + a^{k-m},$$
$$h = a^s + \cdots + a^t.$$

Der Überhang von h über fa^m ist

$$\varrho = \mathrm{Max}(t - k + i - s, t - k, i - s, 0).$$

Man betrachtet zunächst die Lage bei $\varrho = 0$. Falls $t < k$, so gilt

$$Lg\, h a^{-m} = Lg\, h < n = Lg\, f,$$

und man kann (29) anwenden:

$$\boldsymbol{T}\left(\tfrac{1}{h a^{-m}},\ h + fa^m,\ \tfrac{1}{h a^{-m}}(fa^m)\right) = a^m$$
$$\tfrac{1}{f}(fa^m) = a^m$$
$$\tfrac{1}{h a^{-m}} h = a^m$$
$$\boldsymbol{T}\left(\tfrac{1}{f},\ fa^m + h,\ \tfrac{1}{f} h\right) = a^m.$$

Falls aber $t = k$, so hat man nach dem soeben Festgestellten (mit $fa^m + h$ an Stelle von h) die Nebenzerlegbarkeit von

$$\boldsymbol{T}\left(\tfrac{1}{f},\ fa^m + (fa^m + h),\ \tfrac{1}{f}(fa^m + h)\right),$$

aus der man mit (5) die gewünschte Zerlegbarkeit von $\boldsymbol{T}\left(\tfrac{1}{f}, fa^m + h, \tfrac{1}{f} h\right)$ gewinnt.

Dritte Induktionsannahme. (25) sei bei $Lg\, f = n$ schon als richtig erkannt für $\varrho < r\ (> 0)$. Nun soll (25) auch für $\varrho = r$ bewiesen werden. Dazu wendet man (28) mit $\beta = 0$ an.

Fall A ($t < k$) und Fall B ($t = k$):

$$\boldsymbol{T}(a^{m-s}, h + a^s, a^m) = a^{m-s} h$$
$$\boldsymbol{T}(a^{m-s}, fa^m + h, a^m) = a^{m-s}(h + a^s + fa^m)$$
$$\boldsymbol{T}\left(\tfrac{1}{f},\ fa^m + (h + a^s + fa^m), a^m\right) = \tfrac{1}{f}(h + a^s + fa^m)$$
$$\boldsymbol{T}\left(\tfrac{1}{f}, fa^m + h, a^m\right) = \tfrac{1}{f} h.$$

Die beiden ersten Gleichungen sind nach (13) richtig; die dritte Gleichung wird von der dritten Induktionsvoraussetzung erfaßt, weil bei Ersetzung von h durch $h + a^s + fa^m$ der Überhang über fa^m nur von der Änderung am Potenzsummen-Anfang beeinflußt (und zwar erniedrigt) wird.

Fall C ($t > k$):
$$T(a^{m-t}, h + a^t, a^m) = a^{m-t}h$$
$$T(a^{m-t}, fa^m + h, a^m) = a^{m-t}(h + a^t + fa^m)$$
$$T\left(\frac{1}{f}, fa^m + (h + a^t + fa^m), a^m\right) = \frac{1}{f}(h + a^t + fa^m)$$
$$T\left(\frac{1}{f}, fa^m + h, a^m\right) = \frac{1}{f}h.$$

Die beiden ersten Gleichungen sind nach (13) richtig; die dritte fällt unter die dritte Induktionsvoraussetzung, weil bei Ersetzung von h durch $h + a^t + fa^m$ der Überhang über fa^m nur von der Änderung am Potenzsummen-Ende beeinflußt wird.

Damit sind (24) und (25) vollständig bewiesen. Man braucht daher bei dem nun folgenden Beweis von (26) und (27) nur solche Fälle zu betrachten, in denen $Lg\, f$, $Lg\, g > 0$ gilt.

Erste Induktionsannahme. (26) und (27) seien schon als richtig erkannt für alle f, g mit $Lg\, f + Lg\, g < m\ (> 0)$. Nun soll zunächst (26) für $Lg\, f + Lg\, g = m$ hergeleitet werden. Man benutzt wieder die Zerlegung (20). Es sind die beiden Fälle $f_2 \neq 0$ und $f_2 = 0$ zu unterscheiden. — Die folgenden Parallelogramme sind nichtausgeartet, weil die drei uneigentlichen Punkte voneinander verschieden sind und außerdem bei den eigentlichen Punkten verschiedene y-Koordinaten vorkommen.

Fall $f_2 \neq 0$.

I	$U_0\, U_\infty$	U_{1/f_1}	U_{1/f_2}	$U_{1/f}$
II	$y = T\left(\frac{1}{f_1}, x, g^*\right)$	U_{1/f_1}	$(f_1 g^* + f_1 g_1, g_1)$	$(0, g^*)$
III	$y = T\left(\frac{1}{f_1}, x, \frac{1}{f_1}(f_2 g + f_2 g_1 + f_1 g_2)\right)$	U_{1/f_1}	$(f_2 g + f_2 g_1, g_2)$	$(fg_1 + fg_2, g)$
IV	$y = T\left(\frac{1}{f_2}, x, \frac{1}{f_2}(f_1 g^* + f_2 g_1 + f_1 g_1)\right)$	U_{1/f_2}	$(f_1 g^* + f_1 g_1, g_1)$	$(fg_1 + fg_2, g)$
V	$y = T\left(\frac{1}{f_2}, x, g^*\right)$	U_{1/f_2}	$(0, g^*)$	$(f_2 g + f_2 g_1, g_2)$
VI	$y = T\left(\frac{1}{f}, x, \frac{1}{f}(fg_1 + f_1 g^* + f_1 g_1)\right)$	$U_{1/f}$	$(f_1 g^* + f_1 g_1, g_1)$	$(f_2 g + f_2 g_1, g_2)$
VII	$y = T\left(\frac{1}{f}, x, g^*\right)$	$U_{1/f}$	$(0, g^*)$	

Fall $f_2 = 0$.

I	$U_0\, U_\infty$	U_{1/f_1}	U_∞	$U_{1/f}$
II	$y = T\left(\frac{1}{f_1}, x, g^*\right)$	U_{1/f_1}	$(f_1 g^* + f_1 g_1, g_1)$	$(0, g^*)$
III	$y = T\left(\frac{1}{f_1}, x, g_2\right)$	U_{1/f_1}	$(0, g_2)$	$(fg_1 + fg_2, g)$
IV	$x = f_1 g^* + f_1 g_1$	U_∞	$(f_1 g^* + f_1 g_1, g_1)$	$(fg_1 + fg_2, g)$
V	$x = 0$	U_∞	$(0, g^*)$	$(0, g_2)$
VI	$y = T\left(\frac{1}{f}, x, g_2\right)$	$U_{1/f}$	$(f_1 g^* + f_1 g_1, g_1)$	$(0, g_2)$
VII	$y = T\left(\frac{1}{f}, x, g^*\right)$	$U_{1/f}$	$(0, g^*)$	

Bei beiden Parallelogrammen lassen sich die behaupteten Inzidenzen mit der Induktionsvoraussetzung bestätigen, in der letzten Spalte jeweils unter Benutzung der beim Beweis von (18) angegebenen Identitäten. Man bekommt daher gemäß (AF) und nach Anwendung der Identität $g_1+g_2=g+g^*$:

$$T\left(\frac{1}{f}, fg^*+fg, g^*\right)=g.$$

Andererseits ist nach Induktionsvoraussetzung:

$$T\left(\frac{1}{f}, fg^*+fg, g^*\right)=\frac{1}{f}(fg).$$

Folglich gilt

$$\frac{1}{f}(fg)=g.$$

Diese Kürzungsregel wird bei dem anschließenden Beweis von (27), der nun noch für $Lg\,f+Lg\,g=m$ zu erbringen ist, benutzt.

Man verfährt in Anlehnung an den Beweis von (19). Es sei wie dort

$$f=a^i+\cdots+a^j,$$
$$g=a^k+\cdots+a^l,$$
$$h=a^s+\cdots+a^t.$$

[Es kann $h\neq 0$ angenommen werden, da sonst (27′) mit (26) zusammenfällt.]

Als Überhang von h über a^ig ergibt sich

$$\varrho=\mathrm{Max}\left(t-(j+l)+(j+k)-s,\ t-(j+l),\ (j+k)-s,\ 0\right).$$

Bei $\varrho=0$ ist $\quad t-j-l\leq 0$

und $\quad j+k-s\leq 0,$

also $\quad Lg\,h=t-s\leq l-k=Lg\,g.$

Man betrachtet nun, um die Behauptung für $\varrho=0$ zu beweisen, zuerst den Fall $t=j+l$. Dann gilt nach dem eben Bemerkten

$$Lg(g+ha^{-j})<Lg\,g.$$

Ferner ist offenbar

$$(30)\qquad \begin{cases} Lg(f+a^j)<Lg\,f \\ Lg(g+a^l)<Lg\,g. \end{cases}$$

Daher kann man (28) auf Grund der ersten Induktionsvoraussetzung folgendermaßen anwenden, indem man in der dritten Zeile (4) in Verbindung mit der Kürzungsregel benutzt:

$$T\left(\frac{1}{f+a^j}, (f+a^j)\,g + a^j g, g\right)$$
$$= T\left(\frac{1}{f+a^j}, (f+a^j)\,(g+ha^{-j}) + a^j g, g + ha^{-j}\right)$$
$$T\left(\frac{1}{f+a^j}, (f+a^j)\,g + a^j g + h, g\right)$$
$$= T\left(\frac{1}{f+a^j}, (f+a^j)\,(g+ha^{-j}) + a^j g + h, g + ha^{-j}\right)$$
$$T\left(\frac{1}{f}, fg, h\right) = T\left(\frac{1}{f}, f(g+ha^{-j}), g + ha^{-j}\right)$$

$$T\left(\frac{1}{f}, fg + h, g\right) = T\left(\frac{1}{f}, f(g+ha^{-j}) + h, g + ha^{-j}\right) = \frac{1}{f}\,h.$$

Für den Fall $t<j+l$ zieht man die Potenzsumme $h+a^{j+l}$ heran. Sie ordnet sich dem eben Erledigten unter und gewährleistet in der folgenden Anwendung von (28) die Zerlegbarkeit des linken Terms der dritten Zeile, während man an den andern erforderlichen Stellen wegen (30) auf Grund der Induktionsvoraussetzung zerlegen kann.

$$T\left(\frac{1}{f+a^j}, (f+a^j)\,g + a^j g + h + a^{j+l}, g\right)$$
$$= T\left(\frac{1}{f+a^j}, (f+a^j)\,(g+a^l) + a^j g + h + a^{j+l}, g + a^l\right)$$
$$T\left(\frac{1}{f+a^j}, (f+a^j)\,g + a^j g + h, g\right)$$
$$= T\left(\frac{1}{f+a^j}, (f+a^j)\,(g+a^l) + a^j g + h, g + a^l\right)$$
$$T\left(\frac{1}{f}, fg + h + a^{j+l}, g\right) = T\left(\frac{1}{f}, f(g+a^l) + h + a^{j+l}, g + a^l\right)$$

$$T\left(\frac{1}{f}, fg + h, g\right) = T\left(\frac{1}{f}, f(g+a^l) + h, g + a^l\right) = \frac{1}{f}\,h.$$

Demnach ist (27) bei $Lg\,f + Lg\,g = m$ für $\varrho = 0$ richtig.

Zweite Induktionsannahme. (27) sei bei $Lg\,f + Lg\,g = m$ bereits für $\varrho < r\ (>0)$ hergeleitet. Nun soll (27) auch für $\varrho = r$ bewiesen werden. Dazu wird mit der üblichen Fallunterscheidung (28) angewendet.

Fall A und B ($t \leq j+l$). In dem folgenden Gleichungssystem kann man den linken Term der dritten Zeile nach der zweiten Induktionsvoraussetzung zerlegen. Die sonst benötigten Zerlegbarkeiten sind durch die erste Induktionsvoraussetzung oder schon durch (25) gesichert.

$$T\left(\frac{1}{f+a^j}, (f+a^j)\,g + h + a^s, g\right) = T\left(\frac{1}{f+a^j}, (f+a^j)\,a^{s-j} + h + a^s, a^{s-j}\right)$$
$$T\left(\frac{1}{f+a^j}, (f+a^j)\,g + a^j g + h, g\right) = T\left(\frac{1}{f+a^j}, (f+a^j)\,a^{s-j} + a^j g + h, a^{s-j}\right)$$
$$T\left(\frac{1}{f}, fg + a^j g + h + a^s, g\right) = T\left(\frac{1}{f}, fa^{s-j} + a^j g + h + a^s, a^{s-j}\right)$$

$$T\left(\frac{1}{f}, fg + h, g\right) = T\left(\frac{1}{f}, fa^{s-j} + h, a^{s-j}\right) = \frac{1}{f}\,h.$$

Im Fall C ($t>j+l$) läßt sich die Anwendbarkeit von (28) ebenso begründen.

$$T\left(\frac{1}{f+a^j}, (f+a^j)\,g+h+a^t, g\right) = T\left(\frac{1}{f+a^j}, (f+a^j)\,a^{t-j}+h+a^t, a^{t-j}\right)$$

$$T\left(\frac{1}{f+a^j}, (f+a^j)\,g+a^j g+h, g\right) = T\left(\frac{1}{f+a^j}, (f+a^j)\,a^{t-j}+a^j g+h, a^{t-j}\right)$$

$$T\left(\frac{1}{f}, fg+a^j g+h+a^t, g\right) = T\left(\frac{1}{f}, fa^{t-j}+a^j g+h+a^t, a^{t-j}\right)$$

$$T\left(\frac{1}{f}, fg+h, g\right) = T\left(\frac{1}{f}, fa^{t-j}+h, a^{t-j}\right) = \frac{1}{f}h.$$

Damit sind (26) und (27) vollständig bewiesen. Die dazu dualen Formeln, nämlich die Rechtskürzungsregel

$$(fg)\frac{1}{g} = f \tag{31}$$

und die Nebenzerlegbarkeit

$$T\left(fg+h, \frac{1}{g}, f\right) = h\frac{1}{g}, \tag{32}$$

$$T\left(fg+h, \frac{1}{g}, h\frac{1}{g}\right) = f, \tag{32'}$$

lassen sich ganz entsprechend herleiten (und zwar wieder auf dem Weg über die Sonderfälle mit $f=a^\mu$). Man gewinnt zunächst die zu (28) duale Regel aus einem Trapez mit den Parallelen $x=0$, $x=\xi$ und $x=\eta$:

$$\left\{\begin{array}{ll} Aus & \left\{\begin{array}{l} T(\varkappa,\xi,\alpha) = T(\mu,\xi,\beta) \\ T(\lambda,\xi,\alpha) = T(\nu,\xi,\beta) \\ T(\varkappa,\eta,\alpha) = T(\nu,\eta,\beta) \end{array}\right. \\ folgt & T(\lambda,\eta,\alpha) = T(\mu,\eta,\beta). \end{array}\right. \tag{33}$$

(Damit ein nichtausgeartetes Trapez vorlag, waren für die Herleitung die Voraussetzungen $0\neq\xi\neq\eta\neq 0$, $\alpha\neq\beta$ zu machen, die sich nachträglich als überflüssig herausstellen.) — Was nun die beim Beweis von (24) und (26) benutzten, durch ihr Inzidenzschema gegebenen Vierecke betrifft, so lassen sich die jetzt benötigten Vierecke durch folgende Vorschrift daraus gewinnen [nach welcher übrigens schon das Trapez des Beweises von (16) aus dem Parallelogramm des Beweises von (12) gebildet wurde]: Man ersetze einen eigentlichen Punkt (ξ,η) durch eine Gerade $y=T(\xi,x,\eta)$, einen uneigentlichen Punkt $U_\varkappa$ durch eine Gerade $x=\varkappa$, den Punkt U_∞ durch die Gerade U_0U_∞ und umgekehrt. — Diese kurzen Angaben mögen zum Beweis von (31) und (32) genügen.

Die Kürzungsregeln (26) und (31) lassen sich auch in die Form

$$f\left(\frac{1}{f}g\right) = g, \tag{34}$$

$$\left(g\frac{1}{f}\right)f = g \tag{35}$$

bringen. Das geschieht unter Zuhilfenahme der in § 2 erläuterten Sechseckbedingung: Man nimmt mit den Trägerpunkten U_∞, U_0, O für die Linkskürzungsregel das Sechseck

I	$x = \frac{1}{f} g$	$\left(\frac{1}{f} g, g\right)$	$\left(\frac{1}{f} g, \frac{1}{f} g\right)$	
II	$x = g$	(g, fg)	$\left(g, \frac{1}{f} g\right)$	(g, g)
III	$x = fg$	(fg, fg)	(fg, g)	
IV	$y = fg$	(fg, fg)	(g, fg)	
V	$y = g$	(fg, g)	$\left(\frac{1}{f} g, g\right)$	(g, g)
VI	$y = \frac{1}{f} g$	$\left(g, \frac{1}{f} g\right)$	$\left(\frac{1}{f} g, \frac{1}{f} g\right)$	
VII	$y = \frac{1}{f} x$	$\left(g, \frac{1}{f} g\right)$	(fg, g)	
VIII	$y = x$	$\left(\frac{1}{f} g, \frac{1}{f} g\right)$	(fg, fg)	(g, g)
IX	$y = fx$		(g, fg)	

und für die Rechtskürzungsregel das Sechseck

I	$x = f$	(f, g)	(f, gf)	
II	$x = 1$	$\left(1, g \frac{1}{f}\right)$	$(1, gf)$	$(1, g)$
III	$x = \frac{1}{f}$	$\left(\frac{1}{f}, g \frac{1}{f}\right)$	$\left(\frac{1}{f}, g\right)$	
IV	$y = g \frac{1}{f}$	$\left(\frac{1}{f}, g \frac{1}{f}\right)$	$\left(1, g \frac{1}{f}\right)$	
V	$y = g$	$\left(\frac{1}{f}, g\right)$	(f, g)	$(1, g)$
VI	$y = gf$	$(1, gf)$	(f, gf)	
VII	$y = (gf) x$	$(1, gf)$	$\left(\frac{1}{f}, g\right)$	
VIII	$y = gx$	(f, gf)	$\left(\frac{1}{f}, g \frac{1}{f}\right)$	$(1, g)$
IX	$y = \left(g \frac{1}{f}\right) x$		$\left(1, g \frac{1}{f}\right)$	

Die Inzidenz $P_0 \in IX$, die dann folgt, liefert gerade (34) bzw. (35).

Nun hat man nach (26)

$$\frac{1}{g}(gf) = f$$

und nach (35)

$$\left(f \frac{1}{gf}\right)(gf) = f.$$

Daraus folgt

(36)
$$f \frac{1}{gf} = \frac{1}{g}.$$

Entsprechend erhält man aus (31) und (34)

(37) $$\frac{1}{fg}f=\frac{1}{g}.$$

Aus dem folgenden Sechseck mit den Trägerpunkten U_0, U_∞, O kann man die Regel

(38) $$\frac{1}{\frac{1}{f}g}\frac{1}{f}=\frac{1}{g}$$

herleiten; die Bestätigung der ersten Inzidenz in Zeile *VIII* erfolgt wegen $fg=gf$ mit (36).

I	$y=\frac{1}{g}$	$\left(\frac{1}{f},\frac{1}{g}\right)$	$\left(\frac{1}{fg},\frac{1}{g}\right)$	
II	$y=1$	$\left(\frac{1}{f}g,1\right)$	$\left(\frac{1}{fg},1\right)$	$\left(\frac{1}{f},1\right)$
III	$y=g$	$\left(\frac{1}{f}g,g\right)$	$\left(\frac{1}{f},g\right)$	
IV	$x=\frac{1}{f}g$	$\left(\frac{1}{f}g,g\right)$	$\left(\frac{1}{f}g,1\right)$	
V	$x=\frac{1}{f}$	$\left(\frac{1}{f},g\right)$	$\left(\frac{1}{f},\frac{1}{g}\right)$	$\left(\frac{1}{f},1\right)$
VI	$x=\frac{1}{fg}$	$\left(\frac{1}{fg},1\right)$	$\left(\frac{1}{fg},\frac{1}{g}\right)$	
VII	$y=(fg)x$	$\left(\frac{1}{fg},1\right)$	$\left(\frac{1}{f},g\right)$	
VIII	$y=fx$	$\left(\frac{1}{fg},\frac{1}{g}\right)$	$\left(\frac{1}{f}g,g\right)$	$\left(\frac{1}{f},1\right)$
IX	$y=\frac{1}{\frac{1}{f}g}x$		$\left(\frac{1}{f}g,1\right)$	

Man gewinnt damit auch

(39) $$\frac{1}{h}\frac{1}{f}=\frac{1}{fh},$$

denn nach (26) und (38) sind beide Ausdrücke gleich $\frac{1}{\frac{1}{f}(fh)}\frac{1}{f}$. Es lassen sich (36) und (37) daher auch in der Form

$$f\left(\frac{1}{f}\frac{1}{g}\right)=\frac{1}{g}=\left(\frac{1}{g}\frac{1}{f}\right)f$$

schreiben. Mit dem Verfahren, nach dem (34) und (35) gewonnen wurden, kann man dies umwandeln in

(40) $$\frac{1}{f}\left(f\frac{1}{g}\right)=\frac{1}{g}$$

(41) $$\left(\frac{1}{g}f\right)\frac{1}{f}=\frac{1}{g}.$$

(Es sind diesmal Sechsecke mit $P_0=\left(\frac{1}{g},\frac{1}{g}\right)$ bzw. $P_0=\left(1,\frac{1}{g}\right)$ zu benutzen.)

Aus (39) folgt

$$a^k \frac{1}{f} = \frac{1}{f a^{-k}} = \frac{1}{f} a^k. \tag{42}$$

Der Vergleich von (38) mit (41) liefert

$$\frac{1}{\frac{1}{f} g} = \frac{1}{g} f. \tag{43}$$

Es sollen nun noch die dualen Regeln zu (38) und (43) hergeleitet werden. Man betrachtet folgendes Sechseck mit den Trägerpunkten U_∞, U_0, O, bei dem die Größe ω definiert ist durch $\frac{1}{f}\omega = \frac{1}{g}$. Damit bestätigt man die erste Inzidenz in Zeile *VIII;* die erste Inzidenz der Zeile *VII* gilt wegen (37).

I	$x = \omega$	$(\omega, 1)$	$\left(\omega, \frac{1}{g}\right)$	
II	$x = f$	(f, g)	$\left(f, \frac{1}{g}\right)$	$(f, 1)$
III	$x = fg$	(fg, g)	$(fg, 1)$	
IV	$y = g$	(fg, g)	(f, g)	
V	$y = 1$	$(fg, 1)$	$(\omega, 1)$	$(f, 1)$
VI	$y = \frac{1}{g}$	$\left(f, \frac{1}{g}\right)$	$\left(\omega, \frac{1}{g}\right)$	
VII	$y = \frac{1}{fg} x$	$\left(f, \frac{1}{g}\right)$	$(fg, 1)$	
VIII	$y = \frac{1}{f} x$	$\left(\omega, \frac{1}{g}\right)$	(fg, g)	$(f, 1)$
IX	$y = \left(g \frac{1}{f}\right) x$		(f, g)	

Es folgt $\left(g \frac{1}{f}\right)\omega = 1$, also $\omega = \frac{1}{g \frac{1}{f}}$. Den so erhaltenen Wert setzt man in die Definitionsgleichung von ω ein und bekommt

$$\frac{1}{f} \frac{1}{g \frac{1}{f}} = \frac{1}{g}.$$

Der Vergleich mit (40) liefert

$$\frac{1}{g \frac{1}{f}} = f \frac{1}{g}. \tag{44}$$

§ 5. Brüche

Es lassen sich nun neue Hauptzerlegbarkeits-Aussagen machen.

$$T\left(f, \frac{1}{f} g, h\right) = g + h. \tag{45}$$

Beweis. Die Behauptung ist richtig für $Lg\, g = 0$, denn man kann dann mit (42), (31) und (32′) folgendermaßen rechnen:

$$\begin{aligned} \boldsymbol{T}\left(f, \tfrac{1}{f} a^k, h\right) &= \boldsymbol{T}\left(f, \tfrac{1}{f a^{-k}}, h\right) \\ &= \boldsymbol{T}\left((a^k + h) f a^{-k} + h f a^{-k}, \tfrac{1}{f a^{-k}}, (h f a^{-k}) \tfrac{1}{f a^{-k}}\right) \\ &= a^k + h. \end{aligned}$$

Damit hat man aber den Anfang für einen Induktionsbeweis nach dem Muster des Beweises von (13). Man setzt wie dort als erste Induktionsannahme die Richtigkeit von (45) für alle g mit $Lg\, g < n\ (> 0)$ voraus, betrachtet dann $g = a^i + \cdots + a^k$ mit $Lg\, g = n$ und definiert, falls $h = a^s + \cdots + a^t$, als Überhang von h über g

$$\varrho = \mathrm{Max}(t - k + i - s,\ t - k,\ i - s, 0).$$

Bei $\varrho = 0$ ist demgemäß $t - k \leqq 0$ und $i - s \leqq 0$. Falls $t < k$, so gilt $Lg\, h < Lg\, g = n$, und wegen $h = f\left(\frac{1}{f} h\right)$ kann man mittels (5) umwandeln und dann auf Grund der Induktionsvoraussetzung zerlegen. Falls aber $t = k$, so hat man nach dem eben Festgestellten die Zerlegbarkeit von $\boldsymbol{T}\left(f, \frac{1}{f} g, h + g\right)$, aus der man mit (3) die gewünschte Zerlegbarkeit des Terms $\boldsymbol{T}\left(f, \frac{1}{f} g, h\right)$ gewinnt. — Man kann daher als zweite Induktionsannahme die Richtigkeit von (45) bei $Lg\, g = n$ schon für $\varrho < r\ (> 0)$ voraussetzen und muß nun noch (45) für $\varrho = r$ beweisen, was mit der bekannten Fallunterscheidung unter Anwendung von (2) wie folgt geschieht.

Fall A $(t < k)$ und Fall B $(t = k)$:

$$\frac{\boldsymbol{T}\left(f, \frac{1}{f} g, g + h + a^s\right) = \boldsymbol{T}\left(f, \frac{1}{f} a^s, h\right)}{\boldsymbol{T}\left(f, \frac{1}{f} g, h\right) = \boldsymbol{T}\left(f, \frac{1}{f} a^s, g + h + a^s\right) = g + h.}$$

Fall C $(t > k)$:

$$\frac{\boldsymbol{T}\left(f, \frac{1}{f} g, g + h + a^t\right) = \boldsymbol{T}\left(f, \frac{1}{f} a^t, h\right)}{\boldsymbol{T}\left(f, \frac{1}{f} g, h\right) = \boldsymbol{T}\left(f, \frac{1}{f} a^t, g + h + a^t\right) = g + h.}$$

Der linke Term der oberen Zeile fällt hier und vorher unter die zweite Induktionsvoraussetzung.

Damit ist (45) vollständig bewiesen, und unter Verwendung der dualen Formeln läßt sich offenbar genau so auch

$$\boldsymbol{T}\left(f \tfrac{1}{g}, g, h\right) = f + h \tag{46}$$

herleiten:

Zwecks Gewinnung weiterer Rechenregeln betrachtet man jetzt das folgende Sechseck mit den Trägerpunkten U_0, U_∞ und $(0, h)$. Es kann $g \neq 0$ angenommen werden, da sonst die Folgerung (47) nichts Neues liefert.

I	$y = \frac{1}{f} g + h$	$\left(g, \frac{1}{f} g + h\right)$	$\left(\frac{1}{f} g, \frac{1}{f} g + h\right)$	
II	$y = g + h$	$(fg, g + h)$	$\left(\frac{1}{f} g, g + h\right)$	$(g, g + h)$
III	$y = fg + h$	$(fg, fg + h)$	$(g, fg + h)$	
IV	$x = fg$	$(fg, fg + h)$	$(fg, g + h)$	
V	$x = g$	$(g, fg + h)$	$\left(g, \frac{1}{f} g + h\right)$	$(g, g + h)$
VI	$x = \frac{1}{f} g$	$\left(\frac{1}{f} g, g + h\right)$	$\left(\frac{1}{f} g, \frac{1}{f} g + h\right)$	
VII	$y = \boldsymbol{T}(f, x, h)$	$\left(\frac{1}{f} g, g + h\right)$	$(g, fg + h)$	
VIII	$y = x + h$	$\left(\frac{1}{f} g, \frac{1}{f} g + h\right)$	$(fg, fg + h)$	$(g, g + h)$
IX	$y = \boldsymbol{T}\left(\frac{1}{f}, x, h\right)$		$(fg, g + h)$	

Die erste Inzidenz der Zeile *VII* bestätigt man mit (45); für die Inzidenz der letzten Zeile wendet man (25′) an:

$$\boldsymbol{T}\left(\frac{1}{f}, fg, h\right) = \boldsymbol{T}\left(\frac{1}{f}, f(g + h) + fh, \frac{1}{f}(fh)\right) = g + h.$$

Man entnimmt als Folgerung die Inzidenz

$$\boldsymbol{T}\left(\frac{1}{f}, g, h\right) = \frac{1}{f} g + h.$$

Andererseits ist nach (25)

$$\boldsymbol{T}\left(\frac{1}{f}, g, h\right) = \boldsymbol{T}\left(\frac{1}{f}, fh + (fh + g), h\right) = \frac{1}{f}(fh + g),$$

so daß man das Ergebnis bekommt:

$$\frac{1}{f}(fh + g) = h + \frac{1}{f} g. \tag{47}$$

Das nächste Sechseck hat die Trägerpunkte U_0, U_∞ und $(h, 0)$. Es kann $g \neq 0$ angenommen werden, da sonst die zu ziehende Folgerung trivial ist.

I	$y = \frac{1}{f} g$	$\left(g + h, \frac{1}{f} g\right)$	$\left(\frac{1}{f} g + h, \frac{1}{f} g\right)$	
II	$y = g$	$(fg + h, g)$	$\left(\frac{1}{f} g + h, g\right)$	$(g + h, g)$
III	$y = fg$	$(fg + h, fg)$	$(g + h, fg)$	
IV	$x = fg + h$	$(fg + h, fg)$	$(fg + h, g)$	
V	$x = g + h$	$(g + h, fg)$	$\left(g + h, \frac{1}{f} g\right)$	$(g + h, g)$

VI	$x = \frac{1}{f} g + h$	$\left(\frac{1}{f} g + h, g\right)$	$\left(\frac{1}{f} g + h, \frac{1}{f} g\right)$	
VII	$y = T(f, x, fh)$	$\left(\frac{1}{f} g + h, g\right)$	$(g + h, fg)$	
VIII	$y = x + h$	$\left(\frac{1}{f} g + h, \frac{1}{f} g\right)$	$(fg + h, fg)$	$(g + h, g)$
IX	$y = T\left(\frac{1}{f}, x, \frac{1}{f} h\right)$		$(fg + h, g)$	

Die erste Inzidenz der Zeile *VII* bestätigt man mit (47) und (45); für die Inzidenz der Zeile *IX* hat man (27′) zur Verfügung. — Es folgt

$$T\left(\frac{1}{f}, g + h, \frac{1}{f} h\right) = \frac{1}{f} g,$$

oder mit geänderter Bezeichnung

$$T\left(\frac{1}{f}, g, \frac{1}{f} h\right) = \frac{1}{f} (g + h). \tag{48}$$

Um die duale Formel zu gewinnen, betrachtet man zuerst das folgende Sechseck mit den Trägerpunkten U_0, U_∞ und $(0, h)$. Es kann $f \neq 0$ angenommen werden, da sonst die Folgerung (49) nichts Neues liefert.

I	$y = (f + gh) \frac{1}{g}$	$\left(1, (f + gh) \frac{1}{g}\right)$	$\left(\frac{1}{g}, (f + gh) \frac{1}{g}\right)$	
II	$y = f + h$	$(g, f + h)$	$\left(\frac{1}{g}, f + h\right)$	$(1, f + h)$
III	$y = fg + h$	$(g, fg + h)$	$(1, fg + h)$	
IV	$x = g$	$(g, fg + h)$	$(g, f + h)$	
V	$x = 1$	$(1, fg + h)$	$\left(1, (f + gh) \frac{1}{g}\right)$	$(1, f + h)$
VI	$x = \frac{1}{g}$	$\left(\frac{1}{g}, f + h\right)$	$\left(\frac{1}{g}, (f + gh) \frac{1}{g}\right)$	
VII	$y = T(fg, x, h)$	$\left(\frac{1}{g}, f + h\right)$	$(1, fg + h)$	
VIII	$y = T(f, x, h)$	$\left(\frac{1}{g}, (f + gh) \frac{1}{g}\right)$	$(g, fg + h)$	$(1, f + h)$
IX	$y = T\left(f \frac{1}{g}, x, h\right)$		$(g, f + h)$	

Die erste Inzidenz der Zeile *VII* und der Zeile *VIII* bestätigt man mit (32′) bzw. (32):

$$T\left(fg, \frac{1}{g}, h\right) = T\left((f + h) g + hg, \frac{1}{g}, (hg) \frac{1}{g}\right) = f + h,$$

$$T\left(f, \frac{1}{g}, h\right) = T\left(hg + (f + gh), \frac{1}{g}, h\right) = (f + gh) \frac{1}{g}.$$

Für die Inzidenz der Zeile *IX* hat man (46) zur Verfügung. — Es folgt

$$T\left(f \frac{1}{g}, 1, h\right) = (f + gh) \frac{1}{g}.$$

Durch Anwendung von (7) auf den T-Term erhält man daraus

$$T\left(h, 1, f\frac{1}{g}\right) = (f + gh)\frac{1}{g}. \tag{49}$$

Nun wird noch ein Sechseck mit den Trägerpunkten U_h, U_∞ und O ausgewertet. Die im Inzidenzschema auftretende Größe ω wird definiert durch $T\left(h, \frac{1}{g}, \omega\right) = (h+f)\frac{1}{g}$, womit die zweite Inzidenz der ersten Zeile bestätigt ist. — Es kann $f \neq 0$ angenommen werden, da sonst die Folgerung (50) trivial ist.

I	$y = T(h, x, \omega)$	$(1, T(h, 1, \omega))$	$\left(\frac{1}{g}, (f+h)\frac{1}{g}\right)$	
II	$y = T(h, x, f)$	$(g, f + hg)$	$\left(\frac{1}{g}, (fg+h)\frac{1}{g}\right)$	$(1, f+h)$
III	$y = T(h, x, fg)$	$(g, (f+h)g)$	$(1, fg+h)$	
IV	$x = g$	$(g, (f+h)g)$	$(g, f+hg)$	
V	$x = 1$	$(1, fg+h)$	$(1, T(h, 1, \omega))$	$(1, f+h)$
VI	$x = \frac{1}{g}$	$\left(\frac{1}{g}, (fg+h)\frac{1}{g}\right)$	$\left(\frac{1}{g}, (f+h)\frac{1}{g}\right)$	
VII	$y = (fg+h)x$	$\left(\frac{1}{g}, (fg+h)\frac{1}{g}\right)$	$(1, fg+h)$	
VIII	$y = (f+h)x$	$\left(\frac{1}{g}, (f+h)\frac{1}{g}\right)$	$(g, (f+h)g)$	$(1, f+h)$
IX	$y = \left((f+hg)\frac{1}{g}\right)x$		$(g, f+hg)$	

[Hier gehört die Gerade *IX* nicht zum Büschel durch U_h, weil der Faktor von x nach (49) nicht gleich $T(h, 1, 0) = h$ ist.] Die zweite Inzidenz der Zeile *II* bestätigt man mit (32):

$$T\left(h, \frac{1}{g}, f\right) = T\left(fg + (fg+h), \frac{1}{g}, f\right) = (fg+h)\frac{1}{g}.$$

Man entnimmt als Folgerung

$$T(h, 1, \omega) = (f + hg)\frac{1}{g}.$$

Nach (49) ist deshalb $\omega = f\frac{1}{g}$. Setzt man diesen Wert in die Definitionsgleichung von ω ein, so erhält man

$$T\left(h, \frac{1}{g}, f\frac{1}{g}\right) = (f+h)\frac{1}{g}. \tag{50}$$

Damit hat man insbesondere auch die Gleichung

$$T\left(a^k, a^0\frac{1}{g}, h\frac{1}{g}\right) = (a^k + h)\frac{1}{g}$$

zur Verfügung. Sie kann als (erster) Induktionsanfang für einen Beweis der allgemeineren Zerlegbarkeit

$$T\left(a^{k-\mu}, a^\mu\frac{1}{g}, h\frac{1}{g}\right) = (a^k + h)\frac{1}{g} \tag{51}$$

dienen. Muster dafür ist der Beweis von (11). Es ist auch hier nur nötig, positive Exponenten μ zu betrachten. Man setzt als erste Induktionsannahme die Richtigkeit von (51) für $\mu < m\ (>0)$ voraus und hat sie nun für $\mu = m$ nachzuweisen. Der Überhang von h über a^h sei ϱ. Bei $\varrho = 0$ besteht h nur aus der einzigen Potenz a^h; dann ist aber (51) gemäß (4) richtig, weil wegen (42)

$$a^{h-\mu}\left(a^{\mu}\frac{1}{g}\right) = a^{h-\mu}\frac{1}{g\,a^{-\mu}} = \frac{1}{g\,a^{-h}} = a^h\frac{1}{g} \tag{52}$$

gilt. Als zweite Induktionsannahme werde nun bei $\mu = m$ die Richtigkeit von (51) für $\varrho < r\ (>0)$ vorausgesetzt. Der Beweis für $\varrho = r$ läßt sich dann einfach an Hand des Beweises von (11) führen: Man muß dort nur mit (8) an Stelle von (1) arbeiten, weil ja jetzt die a-Potenz mit dem Induktionsexponenten μ als zweites Argument des T-Terms vorkommt, und außerdem an den erforderlichen Stellen den Rechtsfaktor $1/g$ hinzufügen.

Setzt man in (51) $\mu = k$, so ergibt sich

$$T\left(a^0, a^h\frac{1}{g}, h\frac{1}{g}\right) = (a^h + h)\frac{1}{g}.$$

Andererseits ist nach der Definition der Addition

$$T\left(1, a^h\frac{1}{g}, h\frac{1}{g}\right) = a^h\frac{1}{g} + h\frac{1}{g}.$$

Aus der Gleichheit der rechten Seiten entnimmt man aber sofort das allgemeinere Distributivgesetz

$$(f + h)\frac{1}{g} = f\frac{1}{g} + h\frac{1}{g}. \tag{53}$$

In diesem ist als Sonderfall das duale Gegenstück zu (47) enthalten:

$$(fh + g)\frac{1}{f} = h + g\frac{1}{f}. \tag{54}$$

Mit den bisher gewonnenen Regeln soll nun hergeleitet werden

$$\frac{1}{f}g = g\frac{1}{f}. \tag{55}$$

Beweis. (55) ist richtig für $Lg\,f = 0$, sei schon als richtig erkannt für alle f mit $Lg\,f < n\ (>0)$ und soll nun für $Lg\,f = n$ bewiesen werden. Wegen (47) und (54) darf man sich auf den Nachweis von $\frac{1}{f}g' = g'\frac{1}{f}$ mit $Lg\,g' < Lg\,f$ beschränken, denn bei beliebig vorgegebenem f, g läßt sich offenbar immer eine Darstellung $g = fh + g'$ mit $Lg\,g' < Lg\,f$ angeben. Diese eingeschränkte Behauptung erhält man aber mittels (43) und (44) sofort aus der Gleichung

$$\frac{1}{\frac{1}{g'}f} = \frac{1}{f\frac{1}{g'}},$$

deren Richtigkeit von der Induktionsvoraussetzung ausgesagt wird.

Das in (55) stehende Produkt wird von jetzt an *a-Bruch* (oder kürzer *Bruch*) mit dem *Zähler* g und dem *Nenner* f genannt, und es wird dafür die Schreibweise $g/\!/f$ verwendet. Damit ist die bisher schon benutzte Bezeichnung $1/\!/f$ für das Inverse von f verträglich. Zwischen den a-Brüchen und den a-Potenzsummen besteht offensichtlich der Zusammenhang

$$\frac{f}{1} = f. \tag{56}$$

§ 6. Bruchrechnung

In der neuen Schreibweise besitzen die Rechenregeln (43) und (53) die Gestalt

$$\frac{1}{\frac{f}{g}} = \frac{g}{f}, \tag{57}$$

also insbesondere

$$\frac{1}{f a^{-k}} = \frac{a^k}{f}, \tag{58}$$

und

$$\frac{g+h}{f} = \frac{g}{f} + \frac{h}{f}. \tag{59}$$

Die Erweiterungsregel für Brüche

$$\frac{g}{f} = \frac{gh}{fh} \tag{60}$$

läßt sich auf dem schon bei (55) eingeschlagenen Weg beweisen. (60) ist wegen (26) richtig für $Lg\, f = 0$, sei als bereits bewiesen vorausgesetzt für alle f mit $Lg\, f < n\ (> 0)$ und soll nun für $Lg\, f = n$ hergeleitet werden. Dann kann man sich wegen (54) auf die Betrachtung von Brüchen $g'/\!/f$ mit $Lg\, g' < Lg\, f$ beschränken und mit (57) die Behauptung auf die Induktionsvoraussetzung zurückführen:

$$\frac{g'}{f} = \frac{1}{\frac{f}{g'}} = \frac{1}{\frac{fh}{g'h}} = \frac{g'h}{fh}.$$

Nach (59) und (60) ist es also möglich, a-Brüche in der sonst üblichen Weise zu addieren:

$$\frac{f}{p} + \frac{g}{q} = \frac{fq + gp}{pq}. \tag{61}$$

Man ist jetzt in der Lage, die Beweise der §§ 3 und 4 zur Herleitung von Rechenregeln für a-Brüche auszunutzen.

Aus (58) folgt die (schon bekannte) Gleichung

$$a^{\mu} \frac{a^{k-\mu}}{p} = \frac{a^k}{p}.$$

Zusammen mit der Regel (51), die man jetzt in der Form

$$T\left(a^\mu, \frac{a^{k-\mu}}{p}, \frac{h}{p}\right) = \frac{a^k + h}{p}$$

zu schreiben hat, kann sie als Induktionsanfang für den Beweis von

(62) $$a^\mu \frac{g}{p} = \frac{a^\mu g}{p},$$

(63) $$T\left(a^\mu, \frac{g}{p}, \frac{h}{p}\right) = \frac{a^\mu g + h}{p}$$

dienen, der genau dem Beweis von (12), (13) entspricht: Man fügt dort einfach an allen erforderlichen Stellen den Nenner p hinzu.

Wegen (58) und (60) hat man außerdem

$$f \frac{a^\mu}{p} = \frac{f a^\mu}{p}$$

und gewinnt aus (50)

$$T\left(f, \frac{a^\mu}{p}, \frac{h}{p}\right) = \frac{f a^\mu + h}{p}.$$

Diese beiden Gleichungen bilden aber zusammen mit (62), (63) den Induktionsanfang für den Beweis von

(64) $$f \frac{g}{p} = \frac{f g}{p},$$

(65) $$T\left(f, \frac{g}{p}, \frac{h}{p}\right) = \frac{f g + h}{p},$$

der, vom Nenner p abgesehen, fast wörtlich derselbe ist wie der Beweis von (18), (19).

Die dualen Aussagen

(66) $$\frac{f}{p} g = \frac{f g}{p},$$

(67) $$T\left(\frac{f}{p}, g, \frac{h}{p}\right) = \frac{f g + h}{p}$$

lassen sich ohne Schwierigkeit herleiten, da man auf Grund von (58) die Gleichung

$$\frac{a^\mu}{p} a^{k-\mu} = \frac{a^k}{p}$$

und als Sonderfall von (48) unter Berücksichtigung von (58) und (60) die Regel

$$T\left(\frac{a^\mu}{p}, a^{k-\mu}, \frac{h}{p}\right) = \frac{a^k + h}{p}$$

zur Verfügung hat und darauf die erforderlichen Induktionsbeweise aufbauen kann.

Um zu Formeln zu gelangen, in denen ein weiterer Nenner q vorkommt, hat man den Weg, der von dem Gleichungspaar (18), (19) zu den Gln. (64) bis (67) führte, im wesentlichen noch einmal zu durchlaufen.

Man bekommt zunächst mit den Beweisen für (24), (25) und (26), (27):

$$\frac{1}{f}\left(f\frac{a^{\mu}}{p}\right)=\frac{a^{\mu}}{p},$$

$$T\left(\frac{1}{f},\frac{fa^{\mu}+h}{p},\frac{a^{\mu}}{p}\right)=\frac{1}{f}\frac{h}{p},$$

(68)
$$\frac{1}{f}\left(f\frac{g}{p}\right)=\frac{g}{p},$$

(69)
$$T\left(\frac{1}{f},\frac{fg+h}{p},\frac{g}{p}\right)=\frac{1}{f}\frac{h}{p}.$$

[Die erste dieser Gleichungen besitzt man natürlich schon nach (24) und (58).]

Aus (68) erhält man wie bei der Umformung von (26) in (34) unter Heranziehung eines Sechsecks mit $P_0=\left(\frac{g}{p},\frac{g}{p}\right)$:

$$f\left(\frac{1}{f}\frac{g}{p}\right)=\frac{g}{p}.$$

Man hat außerdem nach (64) und (60)

$$f\frac{g}{fp}=\frac{fg}{fp}=\frac{g}{p}.$$

Daher folgt

(70)
$$\frac{1}{f}\frac{g}{p}=\frac{g}{fp}.$$

Auf dem dualen Wege ergeben sich die Gleichungen

$$\left(\frac{f}{p}g\right)\frac{1}{g}=\frac{f}{p},$$

(71)
$$T\left(\frac{fg+h}{p},\frac{1}{g},\frac{f}{p}\right)=\frac{h}{p}\frac{1}{g},$$

(72)
$$\frac{f}{p}\frac{1}{g}=\frac{f}{pg}.$$

Nun kann man weiter aus den in § 5 angegebenen Sechsecken, von denen man jetzt nur das zweite und vierte braucht, Folgerungen ziehen. Wenn man bei dem zweiten Sechseck g und h überall durch g/p und h/p ersetzt und zur Bestätigung der nichttrivialen Inzidenzen (65) und (69) heranzieht, erhält man

$$T\left(\frac{1}{f},\frac{g+h}{p},\frac{h}{fp}\right)=\frac{g}{fp}.$$

Das vierte Sechseck liefert dual dazu, wenn man dort überall f und h durch f/p und h/p ersetzt und die nichttrivialen Inzidenzen mit (67), (71) und (72) bestätigt:

$$T\left(\frac{h}{p},\frac{1}{g},\frac{f}{pg}\right)=\frac{h+f}{pg}.$$

Wegen (58) lassen sich diese beiden Gleichungen mit anderer Bezeichnung auch folgendermaßen schreiben:

$$T\left(\frac{a^\mu}{p}, \frac{g}{q}, \frac{h}{pq}\right) = \frac{a^\mu g + h}{pq},$$

$$T\left(\frac{f}{p}, \frac{a^\mu}{q}, \frac{h}{pq}\right) = \frac{f a^\mu + h}{pq}.$$

Da außerdem nach (70) und (72)

$$\frac{a^\mu}{p}\frac{g}{q} = \frac{a^\mu g}{pq} \quad \text{und} \quad \frac{f}{p}\frac{a^\mu}{q} = \frac{f a^\mu}{pq}$$

gilt, besitzt man die Induktionsanfänge für den Beweis von

(73) $$\frac{f}{p}\frac{g}{q} = \frac{fg}{pq},$$

(74) $$T\left(\frac{f}{p}, \frac{g}{q}, \frac{h}{pq}\right) = \frac{fg + h}{pq},$$

der nach dem Muster des Beweises von (18), (19) verläuft.

(73) ist aber die Regel für die Bruchmultiplikation, die bisher noch fehlte, während durch (74) in Verbindung mit (59), (73) und (60) die Zerlegbarkeit jedes T-Terms ausgesagt wird, dessen Argumente drei beliebige a-Brüche sind.

Die Gln. (56), (60), (61) und (73) besagen, daß der Bereich aller a-Brüche (mit der Addition und Multiplikation des Ternärkörpers) isomorph ist zu dem (formal zu bildenden) Quotientenkörper des Integritätsbereichs aller a-Potenzsummen. Damit ist Satz 3 von der Endlichkeitsvoraussetzung befreit:

Satz 4. *Jeder Anti-*FANO*-Ternärkörper, der von einem Element* $a \neq 0, 1$ *erzeugt wird, ist ein Körper, und zwar eine einfache Erweiterung des Primkörpers der Charakteristik 2.*

Geometrisch bedeutet das:

Satz 5. *In jeder Anti-*FANO*-Ebene, die von einem nichtausgearteten Punktquadrupel und einem weiteren Punkt auf einer Verbindungsgerade erzeugt wird, gilt der Satz von* PAPPOS.

Daraus läßt sich noch folgende Aussage über Schließungssätze, die in allen Anti-FANO-Ebenen richtig sind, gewinnen (wegen der darin vorkommenden Begriffe und wegen des Beweises sei auf [*10*], S. 26ff. und S. 189, verwiesen):

Satz 6. *In jeder Anti-*FANO*-Ebene gelten alle allgemeinen konstruierbaren primitiven Schließungssätze vom Rang 8 und 9, die aus dem Satz von* PAPPOS *folgen.*

Zu diesen Schließungssätzen gehört insbesondere die zweifache Ausartung des Satzes von DESARGUES, die nach [*5*] dem Satz vom vollständigen Vierseit

äquivalent ist. Berücksichtigt man noch das in der Einleitung Gesagte, so ergibt sich:

Satz 7. *Eine notwendige und hinreichende Bedingung dafür, daß in einer projektiven Ebene alle allgemeinen konstruierbaren primitiven Schließungssätze vom Rang 8 und 9 gelten, die aus dem Satz von* PAPPOS *folgen, ist die Gültigkeit der zweifachen Ausartung des Satzes von* DESARGUES.

Ob auch die Gültigkeit der beiden (oder sogar nur eines der beiden) kleinen Sätze von PAPPOS eine solche Bedingung liefert, ist unbekannt. Im Hinblick auf Untersuchungen in dieser Richtung wurde hier mehrfach die Sechseckbedingung herangezogen, obwohl es möglich gewesen wäre, jedes Sechseck durch ein Trapez mit der Eigenschaft (AF) zu ersetzen. [Bei den Sechsecken in § 4 wird das zugehörige Ersatz-Trapez jeweils aus den Punkten P_0, P_1, P_3, P_4, P_6, Q_3 und dem uneigentlichen Punkt der Gerade *VIII* gebildet (vgl. Fig. 2). Dasselbe trifft für die beiden ersten Sechsecke des § 5 zu, und durch Dualisierung der beiden so erhaltenen Trapeze erhält man die Ersatz-Trapeze für die beiden letzten Sechsecke des § 5.]

Literatur

[1] BLASCHKE, W., u. G. BOL: Geometrie der Gewebe, § 4. Berlin 1938. — [2] BOL, G.: Gewebe und Gruppen. Math. Ann. **114**, 414—431 (1937). — [3] HALL, M.: Projective planes. Trans. Amer. Math. Soc. **54**, 229—277 (1943). — [4] MOUFANG, R.: Zur Struktur der projektiven Geometrie der Ebene. Math. Ann. **105**, 536—601 (1931). — [5] MOUFANG, R.: Die Schnittpunktsätze des projektiven speziellen Fünfecknetzes in ihrer Abhängigkeit voneinander. Math. Ann. **106**, 755—795 (1932). — [6] MOUFANG, R.: Ein Satz über die Schnittpunktsätze des allgemeinen Fünfecknetzes. Math. Ann. **107**, 124—139 (1932). — [7] MOUFANG, R.: Die Desarguesschen Sätze vom Rang 10. Math. Ann. **108**, 296—310 (1933). — [8] MOUFANG, R.: Alternativkörper und der Satz vom vollständigen Vierseit. Abh. math. Sem. Univ. Hamburg **9**, 207—222 (1933). — [9] PICKERT, G.: Der Satz vom vollständigen Viereck bei kollinearen Diagonalpunkten. Math. Z. **56**, 131—133 (1952). — [10] PICKERT, G.: Projektive Ebenen. Berlin-Göttingen-Heidelberg 1955.

Tübingen, Mathematisches Institut der Universität

(Eingegangen am 29. November 1955)

VOLKMANN, B.
Math. Zeitschr. Bd. 65, S. 389—413 (1956)

Über Hausdorffsche Dimensionen von Mengen, die durch Zifferneigenschaften charakterisiert sind. V.*)

Von

BODO VOLKMANN

§ 1. Einleitung

1. Ist $g \geq 2$ eine ganze Zahl, so gibt es bekanntlich für jede reelle Zahl ϱ mit $0 < \varrho \leq 1$ eine g-adische Entwicklung oder Zifferndarstellung, d.h. eine Potenzreihe der Form

$$\varrho = \sum_{j=1}^{\infty} \frac{e_j}{g^j}, \tag{1}$$

deren Koeffizienten e_j ganze Zahlen mit $0 \leq e_j < g$ sind. Wir setzen im folgenden stets voraus, daß unendlich viele $e_j \neq 0$ sind, so daß die Darstellung (1) immer eindeutig ist, und nennen die Folge $\{e_j\}$ dann die (g-adische) Zifferfolge von ϱ.

Eine geordnete Menge $\{f_1, \ldots, f_i\}$ von (nicht notwendig verschiedenen) g-adischen Ziffern, abkürzend mit $F = f_1 \ldots f_i$ bezeichnet[1]), nennen wir eine (g-adische) *Sequenz*. Gibt es ein $n \geq 0$, so daß in der Zifferfolge einer Zahl ϱ

$$e_{n+1} \cdots e_{n+i} = f_1 \cdots f_i$$

ist, so sagen wir, F trete bei ϱ auf oder ϱ sei F-haltig; andernfalls nennen wir ϱ eine *F-freie* Zahl. Dieselbe Ausdrucksweise gebrauchen wir auch für Sequenzen E an Stelle unendlicher Zifferfolgen $\{e_j\}$. Als *Länge* $i = |F|$ einer Sequenz F bezeichnen wir die Anzahl ihrer (nicht notwendig verschiedenen) Ziffern.

In einer früheren Arbeit [12][2]) wurde die Menge $K(F; g)$ der F-freien g-adischen Zahlen betrachtet und ihre HAUSDORFFsche Dimension bestimmt. Das dortige Resultat wird in der vorliegenden Arbeit auf die Menge $K(F_1, \ldots, F_m; g)$ der Zahlen ϱ verallgemeinert, bei denen keine Sequenz eines vorgegebenen Systems $\{F_1, \ldots, F_m\}$ auftritt. Für den Fall $m = 1$ wird dabei

---

*) Gekürzte Fassung einer Arbeit, die von der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Mainz als Habilitationsschrift angenommen wurde. — Vorgetragen auf der Tagung über Zahlentheorie im Mathematischen Forschungsinstitut Oberwolfach, 3.—8. Januar 1955.

[1]) Es dürfte keine Gefahr der Verwechslung mit dem Produkt $\prod_{j=1}^{i} f_j$ bestehen.

[2]) Zahlen in eckigen Klammern beziehen sich auf das Literaturverzeichnis am Ende der Arbeit.

ein neuer Beweis erbracht[3]), der über die Dimension α von $K(F; g)$ hinaus noch eine obere und eine untere Schranke für das α-dimensionale HAUSDORFF-sche Maß ergibt.

Als Hilfsmittel verwenden wir bei den nachfolgenden Beweisen die in [12] eingeführten E-Fächer. Einem solchen E-Fächer kann man einerseits eine Matrix $\mathfrak{A}$ und andererseits eine Punktmenge M derart zuordnen, daß die Dimension von M von den Eigenwerten von $\mathfrak{A}$ abhängt. Als solche Mengen M werden sich insbesondere die Mengen $K(F_1, \ldots, F_m; g)$ darstellen lassen. So ergibt sich der für das behandelte Gebiet kennzeichnende Sachverhalt, daß sich hier Zahlentheorie, Maßtheorie und Algebra in eigentümlicher und reizvoller Weise berühren.

2. Hinsichtlich der Terminologie schicken wir folgende Bemerkungen voraus:

1) Ist W eine Menge von Punktmengen, so bezeichnen wir die Vereinigung dieser Punktmengen mit W'.

2) Den Durchmesser einer unendlichen Punktmenge M bezeichnen wir mit $|M|$, ihre abgeschlossene Hülle mit M^*.

3) Ist n eine natürliche Zahl, so bezeichnen wir die g^n Intervalle der Form

$$\nu g^{-n} < \varrho \leq (\nu + 1) g^{-n} \qquad (\nu = 0, \ldots, g^n - 1)$$

als die (g-adischen) *Grundintervalle n-ter Stufe*. Sie treten bei allen metrischen Betrachtungen über Ziffernentwicklungen auf, da die Menge aller Zahlen ϱ, deren g-adische Entwicklung (1) mit einer vorgeschriebenen Ziffernsequenz F der Länge n beginnt, als Punktmenge aufgefaßt, ein Grundintervall n-ter Stufe ist. Wir nennen dann F die zu diesem Grundintervall gehörende Sequenz.

4) Für das Aneinanderfügen zweier Sequenzen $E = e_1 \ldots e_m$ und $F = f_1 \ldots f_n$ verwenden wir die ohne weiteres verständliche Schreibweise $e_1 \ldots e_m f_1 \ldots f_n = EF$. Bezüglich dieser Operation bilden die g-adischen Sequenzen eine Halbgruppe, in der die „leere" Sequenz L, definiert durch die Bedingung

$$EL = LE = E \qquad \text{für jede Sequenz } E,$$

die Rolle eines Einselements spielt. Wir schreiben der leeren Sequenz die Länge 0 zu, so daß immer $|EF| = |E| + |F|$ ist.

5) Ist $E = e_1 \ldots e_n$, so bezeichnen wir die „invertierte" Sequenz $e_n \ldots e_1$ mit $\tilde{E}$.

---

[3]) Ein solcher neuer Beweis war nötig, da der in [12] angegebene leider folgende Lücke enthält: Aus dem zum Beweis von Hilfssatz 3 in [12], S. 265 herangezogenen Ergebnis aus der Arbeit [4] von FROBENIUS folgt nicht, wie dort zitiert, $\gamma_1 > |\gamma_j|$, sondern nur $\gamma_1 \geq |\gamma_j|$. Der Text von [4] ist an der betreffenden Stelle mißverständlich, ebenso die hierauf bezugnehmende Einleitung der nachfolgenden Arbeit „Über Matrizen aus positiven Elementen. II." Sitzgsber. preuß. Akad. Wiss., Physik.-math. Kl. 18, 514—518 (1909).

In [12], Satz 6c) muß es heißen: „... zu $\mathfrak{P}(F')$ oder zu $\mathfrak{P}(F)$" statt „... zu $\mathfrak{P}(F)$ oder zu $\mathfrak{P}(F')$"; ebenso sind in der letzten Ungleichung des Beweises F und F' zu vertauschen.

6) Entsteht F durch Aneinanderfügen von $n \geq 1$ Exemplaren von E, so sagen wir, die Sequenz F sei ein *Vielfaches* der Sequenz E. Dann ist $|F|$ ein Vielfaches von $|E|$.

7) Unter dem *Typ* einer Sequenz E in bezug auf eine Sequenz F, geschrieben $t(E, F)$, verstehen wir die größte Zahl $\varkappa$ derart, daß die Sequenz der letzten $\varkappa$ Ziffern von E der Sequenz der ersten $\varkappa$ Ziffern von F gleicht. Somit ist $t(E, F)$ eine ganze Zahl mit $0 \leq t(E, F) \leq |F|$, und wenn E eine F-freie Sequenz ist, gilt sogar $0 \leq t(E, F) < |F|$.

Kap. 1. Graphentheoretische und maßtheoretische Hilfsbetrachtungen

§ 2. E-Fächer

3. Wie in [12], Nr. 2 bezeichnen wir einen Fächer Ξ dann als *E-Fächer*, wenn es unter seinen (abzählbar unendlich vielen) Unterfächern nur endlich viele verschiedene gibt. Die Anzahl k der verschiedenen Unterfächer eines E-Fächers nennen wir seine *Klasse*. Dann kann Ξ folgendermaßen durch eine k-reihige Matrix $\mathfrak{A}(\Xi)$ beschrieben werden: Sind $\Xi_0, \ldots, \Xi_{k-1}$ die verschiedenen Unterfächer von Ξ und ist etwa $\Xi = \Xi_0$, so sagen wir, ein Knotenpunkt A sei vom *Typ* $\varkappa$, wenn der zugehörige Unterfächer $\Xi(A) = \Xi_\varkappa$ ist. Dann sei das Matrixelement $a_{\varkappa\lambda}(\varkappa, \lambda = 0, \ldots, k-1)$ gleich der Anzahl der Knotenpunkte des Fächers Ξ_λ, die den Index 1 und den Typ $\varkappa$ haben. Durch die Matrix $\mathfrak{A}(\Xi)$ und die Festsetzung $\Xi = \Xi_0$ ist der Fächer Ξ eindeutig bestimmt; denn wenn $\mathfrak{A}(\Xi)$ bekannt ist, läßt sich Ξ sukzessive von seinem Endpunkt A_0 aus konstruieren, da mit jedem Knotenpunkt A, den man bei dieser Konstruktion erhält, vermöge einer Beziehung der Form $\Xi(A) = \Xi_\varkappa$ sein Grad gegeben ist.

4. Die Fächer $\Xi_0, \ldots, \Xi_{k-1}$ sind definitionsgemäß sämtlich E-Fächer, da jeder ihrer Unterfächer einem der $\Xi_\varkappa$ gleicht. Jedoch brauchen für $\varkappa \neq 0$ nicht sämtliche k Fächer als Unterfächer von $\Xi_\varkappa$ aufzutreten, so daß die Klasse von $\Xi_\varkappa$ kleiner als k sein kann. Dennoch gilt, wenn man für jedes ganze $n \geq 0$ die Anzahl der Knotenpunkte des Fächers $\Xi_\varkappa$, deren Index n ist, mit $h_\varkappa(n)$ bezeichnet, der

Satz 1[4]. Die Funktionen $h_\varkappa(n)$ genügen sämtlich der linearen Differenzengleichung mit konstanten Koeffizienten, deren charakteristische Gleichung die charakteristische Gleichung der Matrix $\mathfrak{A}(\Xi)$ ist.

Beweis. Sei $a_{\varkappa\lambda}(n)$ die Anzahl der Knotenpunkte A_n des Fächers Ξ_λ, die den Index n und den zugehörigen Unterfächer $\Xi(A_n) = \Xi_\varkappa$ haben. Dann ist $h_\lambda(n) = \sum_{\varkappa=0}^{k-1} a_{\varkappa\lambda}(n)$. Folglich gilt, wenn man die Elemente der Matrix $\mathfrak{A}(\Xi)^n$ mit $a_{\varkappa\lambda}^{(n)}$ bezeichnet,

$$h_\lambda(n) = \sum_{\varkappa=0}^{k-1} a_{\varkappa\lambda}^{(n)} \quad (n = 0, 1, 2, \ldots;\ \mathfrak{A}(\Xi)^0 = \text{Einheitsmatrix}); \tag{2}$$

[4]) Im wesentlichen von [12], Satz 2 übernommen.

denn es ist

$$a_{\varkappa\lambda}^{(n)} = a_{\varkappa\lambda}(n) \qquad (n = 0, 1, 2, \ldots). \tag{3}$$

Für $n = 0$ gilt die Gl. (3) nämlich trivialerweise, und wenn sie für irgendein n und alle $\varkappa, \lambda$ bewiesen ist, ergibt sich

$$a_{\varkappa\lambda}(n+1) = \sum_{\mu=0}^{k-1} a_{\varkappa\mu} a_{\mu\lambda}(n) = \sum_{\mu=0}^{k-1} a_{\varkappa\mu} a_{\varkappa\lambda}^{(n)} = a_{\varkappa\lambda}^{(n+1)},$$

so daß (3) und damit auch (2) für alle n gilt.

Die charakteristische Gleichung der Matrix $\mathfrak{A}(\Xi)$ sei

$$z^k + q_{k-1} z^{k-1} + \cdots + q_1 z + q_0 = 0.$$

Dann gilt nach der sog. CAYLEYschen Relation[5]) die Matrizengleichung

$$\begin{cases} \mathfrak{A}^k + q_{k-1}\mathfrak{A}^{k-1} + \cdots + q_1\mathfrak{A} + q_0\mathfrak{E} = \mathfrak{N} \\ \mathfrak{E} = \text{Einheitsmatrix}, \quad \mathfrak{N} = \text{Nullmatrix}. \end{cases} \tag{4}$$

Folglich erhält man nach Multiplikation mit $\mathfrak{A}^{n+1-k}$ für $n+1 \geq k$, da die Matrizen auf beiden Seiten von (4) elementweise gleich sind,

$$a_{\varkappa\lambda}^{(n+1)} + q_{k-1} a_{\varkappa\lambda}^{(n)} + \cdots + q_0 a_{\varkappa\lambda}^{(n-k+1)} = 0 \qquad (\varkappa, \lambda = 0, \ldots, k-1).$$

Hieraus ergibt sich nach Summation über $\varkappa$ von 0 bis $k-1$ auf Grund von (2) die behauptete Beziehung

$$h_\lambda(n+1) + q_{k-1} h(n) + \cdots + q_0 h_\lambda(n-k+1) = 0 \qquad (n+1 \geq k).$$

§ 3. Mengen, die einem E-Fächer g-adisch zugeordnet sind

5. Sei Ξ ein E-Fächer, dessen *Grad*[6]) höchstens gleich $g+1$ ist, und sei $h(n) = h_0(n)$ die in § 2 definierte Anzahlfunktion seiner Knotenpunkte. Dann bezeichnen wir eine lineare Punktmenge M als dem Fächer Ξ *(g-adisch) zugeordnet*, falls es eine Folge $I_1, I_2, \ldots$ von Mengen gibt derart, daß folgende Bedingungen erfüllt sind:

a) Die Elemente von I_n sind jeweils $h(n)$ Grundintervalle n-ter Stufe, die mit $\mathfrak{i}_n^1, \ldots, \mathfrak{i}_n^{h(n)}$ bezeichnet werden mögen.

b) Für zwei Intervalle $\mathfrak{i}_n^\nu$ und $\mathfrak{i}_p^\sigma$ gilt genau dann $\mathfrak{i}_n^\nu \subseteq \mathfrak{i}_p^\sigma$, wenn der Knotenpunkt A_n^ν von Ξ dem Unterfächer $\Xi(A_p^\sigma)$ angehört.

c) Die Mengen $M \wedge \mathfrak{i}_n^\nu$ und $M \wedge \mathfrak{i}_p^\sigma$ sind einander im Verhältnis $1 : g^{n-p}$ geometrisch ähnlich, falls die Knotenpunkte A_n^ν und A_p^σ vom gleichen Typ sind.

d) Im Sinne von § 1, Bem. 1) ist

$$M = \bigwedge_{n=0}^{\infty} I_n' \qquad (I_0' = \text{Intervall } 0 < \varrho \leq 1). \tag{5}$$

---

[5]) Siehe z. B. MACDUFFEE [2], S. 18.

[6]) Unter dem Grad eines Fächers verstehen wir das Maximum der Grade seiner Knotenpunkte.

Sind die Bedingungen a) bis d) erfüllt, so bezeichnen wir die zu i'_n gehörende Sequenz mit E'_n und nennen irgend zwei von den vier Dingen „Knotenpunkt A'_n, Unterfächer $\Xi(A'_n)$, Grundintervall i'_n und Sequenz E'_n" einander zugehörig. Aus b) folgt insbesondere

$$I'_n \geqq I'_{n+1} \qquad (n = 0, 1, \ldots). \tag{6}$$

6. Im folgenden soll für solche Mengen M unter der (bei den beabsichtigten Anwendungen stets erfüllten) Zusatzvoraussetzung, daß die Funktion $h(n)$ für alle nichtnegativen ganzzahligen Werte von a und b der Ungleichung

$$h(a+b) \leq h(a)\, h(b) \tag{7}$$

genügt, die HAUSDORFFsche Dimension bestimmt werden. Es gilt der

Satz 2. Sei Ξ ein E-Fächer, dessen Grad $\leq g+1$ ist und dessen Funktion $h(n)$ der Bedingung (7) genügt, und sei M eine ihm g-adisch zugeordnete Punktmenge. Dann existiert $\chi = \lim\limits_{n\to\infty} \frac{1}{n} \log h(n)$, und es ist

$$\dim M = \frac{\chi}{\log g} = \delta. \tag{8}$$

Beweis. a) Zum Nachweis der Existenz des Grenzwertes bemerken wir, daß nach (7) für natürliches w und n

$$h(w\,n) \leq h(n)^w$$

und folglich

$$\frac{\log h(w\,n)}{w\,n} \leq \frac{\log h(n)}{n} \tag{9}$$

gilt. Ist nun $\underline{\lim}\limits_{n\to\infty} \frac{\log h(n)}{n} = \underline{\chi}$ und $\overline{\lim}\limits_{n\to\infty} \frac{\log h(n)}{n} = \bar{\chi}$, so werde n_0 bei gegebenem $\varepsilon > 0$ so gewählt, daß $\left|\frac{\log h(n_0)}{n_0} - \underline{\chi}\right| < \varepsilon$ ist. Daraus folgt dann wegen (9)

$$\frac{\log h(w\,n_0)}{w\,n_0} \leq \frac{\log h(n_0)}{n_0} < \underline{\chi} + \varepsilon \qquad (w = 1, 2, \ldots)$$

und damit

$$\overline{\lim_{w\to\infty}} \frac{\log h(w\,n_0)}{w\,n_0} \leq \underline{\chi} + \varepsilon. \tag{10}$$

Da jeder Knotenpunkt des Fächers Ξ höchstens vom Grad $g+1$, also höchstens mit g Knotenpunkten des nächsthöheren Index durch Wege der Länge 1 verbunden ist, wird

$$h(j+1) \leq g\, h(j) \qquad (j = 0, 1, \ldots).$$

Folglich gilt, wenn bei beliebigem n die Zahl w so gewählt wird, daß

$$w\,n_0 \leq n < (w+1)\,n_0$$

ist, die Beziehung

$$\frac{\log h(n)}{n} < \frac{\log h(w\,n_0 + n_0)}{w\,n_0} \leq \frac{\log\left(g^{n_0} h(w\,n_0)\right)}{w\,n_0} = \frac{\log h(w\,n_0)}{w\,n_0} + O\left(\frac{1}{w}\right).$$

Daraus ergibt sich nach (10)

$$\overline{\chi} = \overline{\lim_{n\to\infty}} \frac{\log h(n)}{n} \leq \overline{\lim_{w\to\infty}} \frac{\log h(w n_0)}{w n_0} \leq \underline{\chi} + \varepsilon,$$

und somit ist

$$\overline{\chi} = \underline{\chi} = \lim_{n\to\infty} \frac{\log h(n)}{n} = \chi.$$

b) Um zu zeigen, daß

$$\dim M \geq \delta \tag{11}$$

ist, berücksichtigen wir, daß die Menge I'_n eine g^{-n}-Überdeckung von M darstellt[7]). Folglich wird für jedes $\alpha > \delta$, da nach Teil a)

$$h(n) = e^{n(\chi+o(1))} = g^{n(\delta+o(1))}$$

ist, das α-dimensionale HAUSDORFFsche Maß von M

$$\{M\}^\alpha = \lim_{n\to\infty} h(n)\, g^{-n\alpha} = \lim_{n\to\infty} g^{n(\delta-\alpha+o(1))} = 0,$$

also $\dim M \leq \alpha$. Daraus folgt (11).

c) Zum Beweis von Beweis von Satz 2 fehlt nun noch der Nachweis von

$$\dim M \leq \delta. \tag{12}$$

Dazu verwenden wir, wenn $I = \bigvee_{n=1}^{\infty} I_n$ gesetzt wird, den

Hilfssatz 1.[8]) Ist $0 < \alpha \leq 1$, so gibt es zu jedem Intervall $(\eta_1; \eta_2) \subseteqq (0; 1)$ eine endliche Menge $V \subseteqq I$ mit

$$(\eta_1; \eta_2) \wedge M \subseteqq V', \tag{13}$$

$$|\mathfrak{i}| < \eta_2 - \eta_1 \quad \text{für jedes } \mathfrak{i} \in V, \tag{14}$$

$$\sum_{\mathfrak{i}\in V} |\mathfrak{i}|^\alpha < (g+2)(\eta_2 - \eta_1)^\alpha. \tag{15}$$

Beweis. Sei n_0 so gewählt, daß

$$g^{-n_0} < \eta_2 - \eta_1 \leq g^{-n_0+1} \tag{16}$$

ist, und sei V die Menge der Intervalle $\mathfrak{i}_{n_0} \in I_{n_0}$, die mit $(\eta_1; \eta_2)$ einen nichtleeren Durchschnitt haben. Dann gilt (13); denn wegen (5) ist $I'_{n_0} \supseteqq M$, und da alle Intervalle von V die Länge g^{-n_0} haben, gilt wegen (16) auch (14). Schließlich ergibt sich die Gültigkeit von (15) aus der Tatsache, daß die Anzahl der Elemente von V definitionsgemäß höchstens gleich

$$\frac{\eta_2 - \eta_1}{g^{-n_0}} + 2 \leq g + 2$$

ist und folglich

$$\sum_{\mathfrak{i}\in V} |\mathfrak{i}|^\alpha \leq (g+2)\, g^{-n_0\alpha} < (g+2)(\eta_2 - \eta_1)^\alpha$$

wird.

---

[7]) In der Bezeichnungsweise von [*11*].

[8]) Im wesentlichen von [*10*], Satz 31 übernommen.

Es werde nun angenommen, für irgendein α sei $\{M\}^\alpha < \infty$. Ist dann $R = \{\mathfrak{r}_1, \mathfrak{r}_2, \ldots\}$ eine Menge von Intervallen derart, daß R' die abgeschlossene Hülle M^* überdeckt und daß

$$\sum_{\mathfrak{r} \in R} |\mathfrak{r}|^\alpha < c_1 \{M\}^\alpha \tag{17}$$

gilt[9]), so gibt es nach dem HEINE-BORELschen Überdeckungssatz eine endliche Untermenge $S = \{\mathfrak{s}_1, \ldots, \mathfrak{s}_l\}$ von R mit $S' \geqq M$, und wenn wir dann auf jedes Intervall $\mathfrak{s}_\lambda$ den Hilfssatz 1 anwenden, so sichert dieser die Existenz einer endlichen Untermenge $T = \{\mathfrak{t}_1, \ldots, \mathfrak{t}_m\}$ von I, für die $T' \geqq M$ und

$$\sum_{\mu=1}^{m} |\mathfrak{t}_\mu|^\alpha < (g+2) \sum_{\lambda=1}^{l} |\mathfrak{s}_\lambda|^\alpha$$

ist. Somit ergibt sich wegen (17)

$$\{M\}^\alpha > \frac{1}{c_1} \sum_{\mathfrak{r} \in R} |\mathfrak{r}|^\alpha \geq \frac{1}{c_1} \sum_{\lambda=1}^{l} |\mathfrak{s}_\lambda|^\alpha \geq c_2 \sum_{\mu=1}^{m} |\mathfrak{t}_\mu|^\alpha.$$

Also gibt es, wenn $M_\varkappa$ eine dem Fächer $\Xi_\varkappa$ g-adisch zugeordnete Menge ist, eine endliche Menge $T_\varkappa \subseteqq I$ mit $T_\varkappa' \geqq M_\varkappa$ und

$$\{M_\varkappa\}^\alpha > c_2(\varkappa) \sum_{\mathfrak{t} \in T_\varkappa} |\mathfrak{t}|^\alpha \qquad (\varkappa = 0, \ldots, k-1;\ M_0 = M). \tag{18}$$

Ist $\mathfrak{i}_n^j$ ein Intervall von I_n und $\varkappa(j) = \varkappa$ sein Typ, so ist die Menge $M \wedge \mathfrak{i}_n^j$ einer solchen Menge $M_\varkappa$ im Verhältnis $1 : g^n$ geometrisch ähnlich, wie aus der Definitionsbedingung c) der g-adischen Zuordnung hervorgeht. Daher gibt es eine Menge $T_{\varkappa n} \subseteqq I$, so daß $T'_{\varkappa n}$ zu einer solchen Menge $T_\varkappa'$ im Verhältnis $1 : g^n$ geometrisch ähnlich ist, ferner $T'_{\varkappa n} \geqq M \wedge \mathfrak{i}_n^j$ und schließlich wegen (18)

$$\{M \wedge \mathfrak{i}_n^j\}^\alpha = g^{-\alpha n} \{M_\varkappa\}^\alpha > c_2(\varkappa)\, g^{-\alpha n} \sum_{\mathfrak{t} \in T_\varkappa} |\mathfrak{t}|^\alpha \tag{19}$$

ist. Setzt man dann

$$c_3 = \min\left(c_2(0), \ldots, c_2(k-1)\right)$$

und $U_n = \bigvee_{j=1}^{k(n)} T_{\varkappa(j),n}$, so ist $U_n \subseteqq \bigvee_{\nu=n}^{\infty} I_\nu$, $U_n' \geqq M$ und wegen (19)

$$\sum_{\mathfrak{t} \in U_n} |\mathfrak{t}|^\alpha < \frac{1}{c_3} \sum_{j=1}^{k(n)} \{M \wedge \mathfrak{i}_n^j\}^\alpha = \frac{1}{c_3} \{M\}^\alpha; \tag{20}$$

letzteres weil nach (5) und (6) offenbar $M = \bigwedge_{\nu=n}^{\infty} I_\nu' \subseteqq I_n'$ und daher $M = \bigvee_{j=1}^{k(n)} (M \wedge \mathfrak{i}_n^j)$ ist. Weiter folgt aus der geometrischen Ähnlichkeit von $T'_{\varkappa(j),n}$ mit $T_\varkappa'$, daß es eine von n unabhängige Zahl $d \geq 0$ gibt, so daß

$$U_n \subseteqq \bigvee_{\nu=n}^{n+d} I_\nu \qquad (n = 1, 2, \ldots)$$

[9]) Mit $c_1, \ldots, c_8$ bezeichnen wir positive Konstanten.

ist. Daher verkleinert sich die linke Seite von (20) höchstens, wenn jedes $T'_{\varkappa(j),n}$ durch eine Menge ersetzt wird, die aus einem einzigen, in $T'_{\varkappa(j),n}$ enthaltenen Intervall von $I_{n+\delta}$ besteht. Aus U_n entsteht dadurch eine Menge W_n mit $W_n \subseteq I_{n+\delta}$ und

$$\sum_{t\in U_n} |t|^\alpha \geq \sum_{t\in W_n} |t|^\alpha \geq h(n)\, g^{-(n+\delta)\alpha} = c_4\, h(n)\, g^{-n\alpha},$$

so daß auf Grund von (20)

$$\{M\}^\alpha > c_5\, h(n)\, g^{-n\alpha}$$

folgt. Da aber nach Teil a)

$$h(n) = g^{n(\delta+o(1))}$$

ist, kann somit die obige Annahme $\{M\}^\alpha < \infty$ nur für $\alpha \geq \delta$ gelten, und daraus folgt (12).

Kap. 2. Zahlenmengen mit ausgeschlossenen Sequenzen

§ 4. Die Mengen $K(F_1, \ldots, F_m; g)$

7. Ist $\mathfrak{F} = \{F_1, \ldots, F_m\}$ eine Menge von m (verschiedenen) g-adischen Sequenzen, so bezeichnen wir mit $K(F_1, \ldots, F_m; g) = K(\mathfrak{F}; g)$ die Menge der reellen Zahlen ϱ mit $0 < \varrho \leq 1$, bei derer g-adischer Entwicklung keine dieser Sequenzen auftritt. Solche ϱ wollen wir $\mathfrak{F}$-*frei* nennen. Zwischen $K(\mathfrak{F}; g)$ und den Mengen $K(F_\mu; g)$ der F_μ-freien Zahlen besteht dann offenbar die Beziehung

$$K(\mathfrak{F}; g) = \bigwedge_{\mu=1}^{m} K(F_\mu; g).$$

Bei den Mengen $\mathfrak{F}$ soll im folgenden stets vorausgesetzt werden, daß alle Sequenzen die gleiche Länge haben. Hinsichtlich der Mengen $K(\mathfrak{F}; g)$ bedeutet das keine Beschränkung der Allgemeinheit; denn wenn

$$\max(|F_1|, \ldots, |F_m|) = i$$

gesetzt wird, kann man jede einzelne Sequenz F_μ durch die Menge der g^{i-i_μ} Sequenzen der Form $G = F_\mu E$ mit $|E| = i - i_\mu$ ersetzen. Bezeichnet man das System der so aus sämtlichen F_μ erhaltenen Sequenzen G nach Streichung etwaiger mehrfach vorkommender mit $\{G_1, \ldots, G_q\} = \mathfrak{G}$, so ist $K(\mathfrak{F}; g) = K(\mathfrak{G}; g)$; denn wenn $\varrho \in K(\mathfrak{F}; g)$ ist, so ist ϱ definitionsgemäß $\mathfrak{F}$-frei und folglich erst recht $\mathfrak{G}$-frei, und wenn ϱ nicht zu $K(\mathfrak{F}; g)$ gehört, also etwa F_μ-haltig ist, so ist ϱ auch G_ν-haltig für geeignetes ν, d.h. ϱ gehört nicht zu $K(\mathfrak{G}; g)$.

In dieser Bezeichnungsweise ist stets $1 \leq m \leq g^i$, da es g^i Sequenzen der Länge i gibt. Wir setzen jedoch im folgenden $m < g^i$ voraus, da sonst $K(\mathfrak{F}; g)$ leer ist.

Ist $I_n = I_n(\mathfrak{F})$ die Menge aller Grundintervalle n-ter Stufe, in denen $\mathfrak{F}$-freie Zahlen liegen, so bezeichnen wir die Elemente von I_n wie in § 3 mit $i_n^1, \ldots, i_n^{h(n)}$, und wir ordnen jedem solchen Intervall i_n^j umkehrbar eindeutig den Knotenpunkt A_n^j vom Index n eines Fächers $\Xi = \Xi(\mathfrak{F})$ zu, der der Bedingung genügen soll, daß genau dann A_n^r ein Knotenpunkt des Fächers

$\Xi(A_p^s)$ ist, wenn $i_n^r \leqq i_p^s$ gilt. Dadurch ist $\Xi(\mathfrak{F})$ eindeutig bestimmt, und es gilt

$$K(\mathfrak{F}; g) = \bigwedge_{n=1}^{\infty} I_n(\mathfrak{F})'.$$

Für irgendeine Sequenz E sei[10])

$$\varkappa = \max_{\mu=1,\ldots,m} t(E, F_\mu),$$

und ν sei der kleinste Wert von μ, für den dieses Maximum angenommen wird. Dann bezeichnen wir das geordnete Zahlenpaar $(\varkappa, \nu)$ als den Typ von E in bezug auf $\mathfrak{F}$, geschrieben $t(E, \mathfrak{F}) = (\varkappa, \nu)$. Sind $\Xi(A_n^r)$ und $\Xi(A_p^s)$ zwei Unterfächer von $\Xi(\mathfrak{F})$, für die in diesem Sinne

$$t(E_n^r, \mathfrak{F}) = t(E_p^s, \mathfrak{F})$$

ist, so gilt für jede beliebige Sequenz D der

Hilfssatz 2. Ist von den beiden Sequenzen $E_n^r D$ und $E_p^s D$ die eine $\mathfrak{F}$-frei, so auch die andere.

Beweis. Ist D nicht $\mathfrak{F}$-frei, so ist die Behauptung trivial. Bei $\mathfrak{F}$-freiem D ist, falls etwa $E_n^r D$ nicht $\mathfrak{F}$-frei, sondern z.B. F_μ-haltig ist, $E_n^r D$ von der Form $E_n^r D = A F_\mu B$. Dabei müssen die Ziffern von F_μ teils zu E_n^r und teils zu D gehören, da ja sowohl E_n^r als auch D nach Voraussetzung F_μ-frei sind. Somit ist $t(E_n^r, F_\mu) > 0$. Da aber

$$t(E_n^r, \mathfrak{F}) = t(E_p^s, \mathfrak{F})$$

und folglich auch $t(E_n^r, F_\mu) = t(E_p^s, F_\mu)$ ist, tritt F_μ auch bei $E_p^s D$ auf, nämlich in einem Exemplar, das sich aus den letzten $t(E_p^s, F_\mu)$ Ziffern von E_p^s und den ersten $|F_\mu| - t(E_p^s, F_\mu)$ Ziffern von D zusammensetzt.

Aus Hilfssatz 2 ist leicht zu folgern, daß die Unterfächer $\Xi(A_n^r)$ und $\Xi(A_p^s)$ gleich sind, falls die Sequenzen E_n^r und E_p^s vom gleichen Typ in bezug auf $\mathfrak{F}$ sind. Also ist $\Xi(\mathfrak{F})$ ein E-Fächer, und die Typengleichheit bei seinen Unterfächern im Sinne von § 2 entspricht der Typengleichheit in bezug auf $\mathfrak{F}$ bei den Sequenzen, die zu den Knotenpunkten von $\Xi(\mathfrak{F})$ gehören.

Die Klasse k von $\Xi(\mathfrak{F})$ ist somit höchstens gleich der Anzahl der verschiedenen Werte, die $t(E, \mathfrak{F})$ annimmt, wenn E alle $\mathfrak{F}$-freien Sequenzen durchläuft. Da dann, wenn wieder die Länge der F_μ mit i bezeichnet wird und $t(E, \mathfrak{F}) = (\varkappa, \nu)$ gesetzt wird, $0 \leq \varkappa \leq i-1$ und $1 \leq \nu \leq m$ ist, im Fall $\varkappa = 0$ aber nur der Wert $\nu = 1$ auftreten kann, wird somit

$$k \leq (i-1)\, m + 1. \tag{21}$$

8. Diese Bemerkungen ermöglichen den Beweis von

Satz 3. Sei $h(\mathfrak{F}; n)$ die Anzahl der $\mathfrak{F}$-freien Sequenzen der Länge n. Dann existiert

$$\lim_{n\to\infty} h(\mathfrak{F}; n)^{1/n} = \gamma(\mathfrak{F}),$$

und es ist

$$\dim K(\mathfrak{F}; g) = \frac{\log \gamma(\mathfrak{F})}{\log g}.$$

[10]) Vgl. § 1, Bem. 7.

Beweis. Die Menge $K(\mathfrak{F};g)=\bigwedge_{n=0}^{\infty} I_n(\mathfrak{F})'$ ist dem E-Fächer $\Xi(\mathfrak{F})$ offenbar g-adisch zugeordnet im Sinne von § 3; denn die dortigen Definitionsbedingungen a), b), d) sind nach den obigen Überlegungen erfüllt, und die Gültigkeit von c) folgt unmittelbar aus der Tatsache, daß für irgend zwei Grundintervalle i_n^{ν} und $i_p^{\varkappa}$ mit den linken Endpunkten ωg^{-n} und λg^{-p} die Mengen $K(\mathfrak{F};g)\wedge i_n^{\nu}$ und $K(\mathfrak{F};g)\wedge i_p^{\varkappa}$ durch die Transformation

$$\varrho'=(\varrho-\omega g^{-n})\,g^{n-p}+\lambda g^{-p}$$

bzw. ihre Umkehrung ineinander übergehen, wenn die zugehörigen Sequenzen vom gleichen Typ sind. Somit ist der Satz 2 mit $h(n)=h(\mathfrak{F};n)$ anwendbar; denn die Ungleichung (7) ist erfüllt, da es für jede $\mathfrak{F}$-freie Sequenz S mit $|S|=a+b$ eine Zerlegung der Form $S=EG$ mit $|E|=a$, $|G|=b$ gibt, bei der E und G dann $\mathfrak{F}$-frei sind. Somit erhält man die Behauptung, wenn man in Satz 2 $e^{\varkappa}=\gamma(\mathfrak{F})$ setzt.

9. Betrachten wir nun die k-reihige Matrix $\mathfrak{A}=\mathfrak{A}(\Xi(\mathfrak{F}))$, die dem E-Fächer $\Xi(\mathfrak{F})$ im Sinne von § 2 nach (21) zugeordnet werden kann. Sind $\gamma_1,\ldots,\gamma_q$ $(q\leq k)$ ihre (reellen oder komplexen) Eigenwerte und $\nu_1,\ldots,\nu_q$ deren Vielfachheiten $\left(\sum_{\varkappa=1}^{q}\nu_{\varkappa}=k\right)$, so folgt aus Satz 1 auf Grund der Theorie der linearen Differenzengleichungen[11]), daß die Funktion $h(\mathfrak{F};n)$ sich in der Form

$$h(\mathfrak{F};n)=P_1(n)\,\gamma_1^n+\cdots+P_q(n)\,\gamma_q^n \qquad (n=0,1,\ldots) \tag{22}$$

darstellen läßt, wobei jedes $P_{\varkappa}(n)$ ein Polynom von einem Grade $\leq\nu_{\varkappa}-1$ ist. Bezüglich dieser Darstellung gilt der

Satz 4. Wenn es unter den Eigenwerten $\gamma_{\varkappa}$ einen (einfachen oder mehrfachen) gibt, etwa γ_1, der positiv-reell und von größerem Betrag ist als alle übrigen, und wenn ferner das Polynom $P_1(n)$ nicht identisch verschwindet, ist $\gamma(\mathfrak{F})=\gamma_1$ und

$$\dim K(\mathfrak{F};g)=\frac{\log\gamma_1}{\log g}. \tag{23}$$

Beweis. Nach (22) gilt, wenn etwa $P_1(n)=\sigma_j n^j+\cdots+\sigma_0 n^0$ $(j\geq 0,\ \sigma_j\neq 0)$ ist, die asymptotische Gleichung

$$h(\mathfrak{F};n)\simeq\gamma_1^n\left(\sigma_j n^j+P_2(n)\left(\frac{\gamma_2}{\gamma_1}\right)^n+\cdots+P_q(n)\left(\frac{\gamma_q}{\gamma_1}\right)^n\right)\simeq\sigma_j n^j\gamma_1^n,$$

woraus wegen $h(\mathfrak{F};n)\to+\infty$ offenbar $\sigma_j>0$ und daher

$$\gamma(\mathfrak{F})=\lim_{n\to\infty}h(\mathfrak{F};n)^{1/n}=\gamma_1$$

folgt. Die Gl. (23) ergibt sich somit aus Satz 3.

Als einfacher Spezialfall von Satz 4 erweist sich der [auch direkt zu beweisende und bereits bekannte[12])]

[11]) Siehe z. B. NÖRLUND [9], S. 297.

[12]) Soweit der Verfasser in Erfahrung bringen konnte, ist dieser Satz zwar nirgends in der Literatur explizit formuliert, doch war er mindestens 1941 bereits Herrn I. J. GOOD [5] bekannt.

Satz 5. Ist $m \leq g$ und ist $\mathfrak{F} = \{f_1, \ldots, f_m\}$ eine Menge[13]) von verschiedenen g-adischen Ziffern, so gilt

$$\dim K(\mathfrak{F}; g) = \frac{\log(g-m)}{\log g}.$$

Beweis. Offenbar ist $h(\mathfrak{F}; n) = (g-m)^n$, also $\gamma_1 = g - m$ im Sinne von Satz 4.

10. Bezeichnet man, wenn $\mathfrak{F} = \{F_1, \ldots, F_m\}$ eine Menge von m Sequenzen der gleichen Länge i ist, mit $\mathfrak{P}_{\mu\nu}(\mathfrak{F}) = \mathfrak{P}_{\mu\nu} = \{p_{\mu\nu 1}, p_{\mu\nu 2}, \ldots, p_{\mu\nu c_{\mu\nu}}\}$ die (möglicherweise leere) Menge der Zahlen p, für die $0 < p < i$ und die Sequenz der ersten p Ziffern von F_μ der der letzten p Ziffern von F_ν gleich ist, so gilt der

Satz 6. Stimmen zwei Sequenzenmengen $\mathfrak{F} = \{F_1, \ldots, F_m\}$ und $\mathfrak{G} = \{G_1, \ldots, G_m\}$ in der Länge i und in den m^2 Mengen $\mathfrak{P}_{\mu\nu}$ $(\mu, \nu = 1, \ldots, m)$ überein, so ist $\dim K(\mathfrak{F}; g) = \dim K(\mathfrak{G}; g)$.

Beweis. Es ist zu zeigen, daß $\gamma(\mathfrak{F})$ nur von der Länge i und den Mengen $\mathfrak{P}_{\mu\nu}$ abhängt. Dazu setzen wir voraus, daß die Elemente $p_{\mu\nu\lambda} \in \mathfrak{P}_{\mu\nu}$ $(\lambda = 1, \ldots, c_{\mu\nu})$ der Größe nach numeriert sind, und betrachten die Vereinigungen

$$\mathfrak{P}_\mu = \bigvee_{\nu=1}^{m} \mathfrak{P}_{\mu\nu} = \{p_{\mu 1}, \ldots, p_{\mu c_\mu}\} \qquad (\mu = 1, \ldots, m),$$

$$\mathfrak{P} = \bigvee_{\mu=1}^{m} \mathfrak{P}_\mu = \{p_1, \ldots, p_c\},$$

in denen ebenfalls die Elemente jeweils der Größe nach numeriert sein sollen. Ferner sei $\mathfrak{N}_{\mu\lambda}$ die Menge der ν mit $p_{\mu\lambda} \in \mathfrak{P}_{\mu\nu}$ und $x_\mu(\mathfrak{F}; j, n) = x_\mu(j, n)$ die Anzahl der $\mathfrak{F}$-freien Sequenzen E mit $|E| = n$ und $t(E, F_\mu) = j$ $(\mu = 1, \ldots, m)$. Dann gilt offenbar für die Funktion $h(\mathfrak{F}; n) = h(n)$ die Gleichung

$$h(n+1) = g\,h(n) - \sum_{\mu=1}^{m} x_\mu(i-1, n) \qquad (n \geq i-1), \tag{24}$$

weil aus einer $\mathfrak{F}$-freien Sequenz E der Länge n genau dann durch Anfügen einer geeigneten Ziffer eine F_μ-haltige Sequenz entsteht, wenn $t(E, F_\mu) = i-1$ ist. Wir unterscheiden nun zwei Fälle:

a) Ist $\mathfrak{P}_\mu$ leer, so wird

$$x_\mu(i-1, n) = h(n-i+1) \qquad (n \geq i-1), \tag{25}$$

da dann einerseits für jede der $h(n-i+1)$ vorhandenen $\mathfrak{F}$-freien Sequenzen E mit $|E| = n-i+1$, wenn $F_\mu = f_1^\mu \ldots f_i^\mu$ ist, offenbar $E f_1^\mu \ldots f_{i-1}^\mu$ eine $\mathfrak{F}$-freie Sequenz der Länge n ist, also $h(n-i+1) \leq x_\mu(i-1, n)$ gilt, und andererseits stets $x_\mu(i-1, n) \leq h(n-i+1)$ ist. Insbesondere folgt für den Fall, daß $\mathfrak{P}$ leer ist, also (25) für jedes μ gilt, aus (24) die Differenzengleichung

$$h(n+1) = g\,h(n) - m\,h(n-i+1) \qquad (n \geq i-1). \tag{26}$$

[13]) Nicht zu verwechseln mit der Sequenz $F = f_1 \ldots f_m$.

b) Ist $\mathfrak{P}_\mu$ nicht leer, so braucht eine Sequenz der Form $E f_1^\mu \dots f_{i-1}^\mu$ mit $\mathfrak{F}$-freiem E nicht $\mathfrak{F}$-frei zu sein, sondern sie kann, wenn $\mathfrak{P}_{\mu\nu}$ nicht leer ist, eine Darstellung der Form

$$E f_1^\mu \dots f_{i-1}^\mu = B F_\nu f_{i-p+1}^\mu \dots f_{i-1}^\mu$$

besitzen, wobei dann $p \in \mathfrak{P}_{\mu\nu}$ ist. Daher erhält man statt (25)

$$(27) \quad x_\mu(i-1,n) = x_\mu(p_{\mu c_\mu} - 1, n-i+p_{\mu c_\mu}) - \sum_{\nu \in \mathfrak{R}_{\mu c_\mu}} x_\nu(i-1, n-i+p_{\mu c_\mu})$$

und entsprechend für $\lambda = 2, \dots, c_\mu$

$$(28) \quad \begin{cases} x_\mu(p_{\mu\lambda} - 1, n-i+p_{\mu\lambda}) = x_\mu(p_{\mu,\lambda-1} - 1, n-i+p_{\mu,\lambda-1}) - \\ \quad - \sum_{\nu \in \mathfrak{R}_{\mu,\lambda-1}} x_\nu(i-1, n-i+p_{\mu,\lambda-1}) \end{cases}$$

sowie schließlich für $\lambda = 1$

$$(29) \quad x_\mu(p_{\mu 1} - 1, n-i+p_{\mu 1}) = h(n-i+1).$$

Setzt man die Gln. (28) und die Gl. (29) nacheinander in (27) ein, so erhält man

$$(30) \quad \begin{cases} x_\mu(i-1,n) = h(n-i+1) - \sum_{\lambda=1}^{c_\mu} \sum_{\nu \in \mathfrak{R}_{\mu\lambda}} x_\nu(i-1, n-i+p_{\mu\lambda}) \\ \qquad (\mu = 1, \dots, m;\ n \geq i-1), \end{cases}$$

also nach (24)

$$(31) \quad \begin{cases} h(n+1) = g\,h(n) - m\,h(n-i+1) + \sum_{\mu=1}^{m} \sum_{\lambda=1}^{c_\mu} \sum_{\nu \in \mathfrak{R}_{\mu\lambda}} x_\nu(i-1, n-i+p_{\mu\lambda}) \\ \qquad (n \geq i-1). \end{cases}$$

Für $0 \leq j \leq i-1$ ist trivialerweise

$$(32) \quad h(\mathfrak{F}; j) = h(\mathfrak{G}; j) \quad \text{und} \quad x_\mu(\mathfrak{F}; i-1; n) = x_\mu(\mathfrak{G}; i-1; n) \ (\mu = 1, \dots, m),$$

da ja $h(j) = g^j$ und

$$x_\mu(i-1; j) = \begin{cases} 0 \text{ für } 0 \leq j < i-1 \\ 1 \text{ für } j = i-1 \end{cases} \qquad (\mu = 1, \dots, m)$$

ist. Wenn aber die Gln. (32) für $j = 0, \dots, n$ mit $n \geq i-1$ bewiesen sind, so folgt aus (31) ihre Gültigkeit für $j = n+1$, so daß für alle $n \geq 1$, wie man entsprechend auch im Fall a) an Hand von (26) erkennt, $h(\mathfrak{F}; n) = h(\mathfrak{G}; n)$ ist. Daraus folgt die Behauptung nach Satz 3.

Übrigens kann man, indem man aus den $m+1$ Gln. (30) und (24) die m Funktionen $x_\mu(n)$ sukzessive eliminiert, eine Differenzengleichung gewinnen, in der nur noch die Funktion $h(n)$ auftritt: Verwendet man lineare Differenzenoperatoren $Q_{\mu\nu}$ der Form

$$Q_{\mu\nu} f(n) = \sum_{j=1}^{i-1} \alpha_{j\mu\nu} f(n-j)$$

und schreibt man abkürzend $x_\mu(i-1, n) = x_\mu(n)$, so lassen sich die Gln. (30) zusammen als ein lineares Gleichungssystem der Form

$$\begin{aligned} Q_{11}\, x_1(n) + \cdots + Q_{1m}\, x_m(n) &= h(n-i+1) \\ \vdots \qquad\qquad\qquad \vdots \qquad & \quad \vdots \\ Q_{m1}\, x_1(n) + \cdots + Q_{mm}\, x_m(n) &= h(n-i+1) \end{aligned}$$

darstellen. Da solche Operatoren bei Addition und Multiplikation einen Ring bilden, kann man auf dieses Gleichungssystem die Sätze der linearen Algebra anwenden. So erhält man, wenn

$$\begin{vmatrix} Q_{11} \cdots Q_{1m} \\ \vdots \qquad \vdots \\ Q_{m1} \cdots Q_{mm} \end{vmatrix} = Q$$

gesetzt und die mit dem Vorzeichen $(-1)^{\mu+\nu}$ versehene $(m-1)$-reihige Unterdeterminante des Elementes $Q_{\mu\nu}$ mit $Q^*_{\mu\nu}$ bezeichnet wird, die Gleichungen

$$Q\, x_\mu(n) = Q^*_\mu\, h(n-i+1) \qquad (\mu = 1, \ldots, m),$$

wobei

$$Q^*_\mu = Q^*_{1\mu} + \cdots + Q^*_{m\mu}$$

ist. Damit ergibt sich nach (24) die gesuchte Differenzengleichung in der Form

$$Q\Big(\sum_{\mu=1}^{m} x_\mu(n)\Big) = \sum_{\mu=1}^{m} Q^*_\mu\, h(n-i+1) = g\, h(n) - h(n+1), \tag{33}$$

d.h.

$$h(n+1) - g\, h(n) - \sum_{\mu=1}^{m} \sum_{\nu=1}^{n} Q^*_{\mu\nu}\, h(n-i+1) = 0. \tag{34}$$

Als unmittelbare Folgerungen erhält man aus dem Beweis von Satz 6:

Folgerung 1. Sei $\Pi = \begin{pmatrix} 0 & 1 & \ldots & g-1 \\ \nu_0 & \nu_1 & & \nu_{g-1} \end{pmatrix}$ irgendeine Permutation der g-adischen Ziffern, die somit jede Sequenz $F = f_1 \ldots f_i$ in eine Sequenz $\Pi(F) = \nu_{f_1} \ldots \nu_{f_i}$ überführt. Dann ist

$$\dim K(F_1, \ldots, F_m; g) = \dim K(\Pi(F_1), \ldots, \Pi(F_m); g).$$

Folgerung 2. Wird für eine Sequenzenmenge $\mathfrak{F} = \{F_1, \ldots, F_m\}$ im Sinne von § 1, Bem. 6) $\tilde{\mathfrak{F}} = \{\tilde{F}_1, \ldots, \tilde{F}_m\}$ gesetzt, so ist

$$\dim K(\mathfrak{F}; g) = \dim K(\tilde{\mathfrak{F}}; g).$$

Denn es ist

$$\mathfrak{P}_{\mu\nu}(\tilde{\mathfrak{F}}) = \mathfrak{P}_{\nu\mu}(\mathfrak{F}) \qquad (\mu, \nu = 1, \ldots, m),$$

folglich auch $Q_{\mu\nu}(\tilde{\mathfrak{F}}) = Q_{\nu\mu}(\mathfrak{F})$ und $Q^*_{\mu\nu}(\tilde{\mathfrak{F}}) = Q^*_{\nu\mu}(\mathfrak{F})$, so daß die Differenzengleichung (34) für $h(\tilde{\mathfrak{F}}; n)$ dieselbe ist wie für $h(\mathfrak{F}; n)$. Da offenbar beide Funktionen für $n = 0, \ldots, i-1$ übereinstimmen, ergibt sich die Behauptung nach Satz 3.

Folgerung 3. Die Folgerungen 1 und 2 können kombiniert werden.

Folgerung 4. Sei $\mathfrak{F}=\{F_1, \ldots, F_m\}$ und $C(\mathfrak{F}; g)$ die Menge der Zahlen ϱ, in deren g-adischer Ziffernfolge jede der Sequenzen F_μ nur endlich oft auftritt. Dann ist

$$\dim C(\mathfrak{F}; g) = \dim K(\mathfrak{F}; g).$$

Denn ist $\varrho \in C(\mathfrak{F}; g)$, so ist ϱ von einer gewissen Stelle an $\mathfrak{F}$-frei, also von der Form

$$\varrho = k g^{-n} + \varrho', \quad k, n \text{ nichtnegativ}, \quad \varrho' \in K(\mathfrak{F}; g),$$

und somit ist $C(\mathfrak{F}; g)$ in einer Vereinigung von abzählbar vielen Mengen enthalten, die sämtlich aus $K(\mathfrak{F}; g)$ durch Translation hervorgehen und daher dieselbe Dimension wie $K(\mathfrak{F}; g)$ haben. Daraus folgt die Behauptung.

11. Im folgenden soll die Frage beantwortet werden, für welche $\mathfrak{F}$ die Dimension von $K(\mathfrak{F}; g)$ verschwindet. Dazu führen wir einige weitere Begriffe ein:

Ist S eine gegebene Sequenz, so bezeichnen wir die Menge aller $|S|$ Sequenzen, die aus S durch zyklische Vertauschung der Ziffern entstehen, als *Zyklus* $\mathfrak{Z}(S)$.

Ein Zyklus heiße $\mathfrak{F}$-frei, wenn jedes seiner Elemente $\mathfrak{F}$-frei ist. Dann gilt der

Hilfssatz 3. Ist S eine Sequenz mit $|S| \geq i$ und P die Sequenz der ersten $i-1$ Ziffern von S, so ist der Zyklus $\mathfrak{Z}(S)$ genau dann $\mathfrak{F}$-frei, wenn die Sequenz SP es ist.

Beweis. Für jede Sequenz $S_1 \in \mathfrak{Z}(S)$ gibt es eine Darstellung der Form $SS = E_1 S_1 E_2$. Folglich gibt es, falls S_1 nicht $\mathfrak{F}$-frei, sondern z.B. F_μ-haltig ist, eine Gleichung der Form

$$SS = E_3 F_\mu E_4. \tag{35}$$

Gehören dabei die Ziffern von F_μ ganz zum ersten oder ganz zum zweiten Exemplar von S, so ist S und damit auch SP nicht $\mathfrak{F}$-frei. Letzteres trifft aber wegen $|P| = |F_\mu| - 1$ auch in dem übrigbleibenden Fall zu, wo die Ziffern von F_μ sich auf beide Exemplare von S verteilen. Daraus folgt die Behauptung; denn umgekehrt ist evident, daß es bei nicht $\mathfrak{F}$-freiem SP eine nicht $\mathfrak{F}$-freie Sequenz $S_1 \in \mathfrak{Z}(S)$ gibt.

Zwei Zyklen $\mathfrak{Z}(S)$ und $\mathfrak{Z}(T)$ nennen wir *wesentlich verschieden*, wenn es weder eine Sequenz von $\mathfrak{Z}(T)$ gibt, die ein Vielfaches[14]) von S ist, noch umgekehrt. In dieser Bezeichnungsweise gilt der

Satz 7. Dann und nur dann ist $\dim K(\mathfrak{F}; g) = 0$, wenn es keine zwei Sequenzen S und T mit $|S| \geq i$, $|T| \geq i$ gibt, die in ihren ersten $i-1$ Ziffern übereinstimmen, derart, daß die Zyklen $\mathfrak{Z}(S)$ und $\mathfrak{Z}(T)$ wesentlich verschieden und beide $\mathfrak{F}$-frei sind. Falls die Dimension von $K(\mathfrak{F}; g)$ verschwindet, ist diese Menge höchstens abzählbar.

[14]) Vgl. § 1, Bem. 5).

Beweis. a) Angenommen, es gibt Sequenzen S und T mit den geforderten Eigenschaften. Dann gibt es Sequenzen P, A, B mit

$$S = PA, \quad T = PB \quad \text{und} \quad |P| = i - 1.$$

Daher sind die Sequenzen $C = ST$ und $D = TS$ wegen $C = PAPB$ und $D = PBPA$ beide $\mathfrak{F}$-frei, wie aus Hilfssatz 3 folgt, da ja die Zyklen $\mathfrak{Z}(PA)$ und $\mathfrak{Z}(PB)$ nach Voraussetzung $\mathfrak{F}$-frei sind. Ist $|S| = q$, $|T| = r$, also $|C| = |D| = q + r$, so entsprechen die Sequenzen C und D im g^{q+r}-adischen System je einer einzigen Ziffer, die wir wieder mit C bzw. D bezeichnen. Ist dann K die Menge aller ϱ, in deren g^{q+r}-adischer Entwicklung ausschließlich die beiden Ziffern C und D auftreten, so gilt $K \subseteq K(\mathfrak{F}; g)$, wie sich bei der Umformung der g^{q+r}-adischen in g-adische Entwicklungen auf Grund der Tatsache ergibt, daß die Sequenzen CC, CD, DC, DD sämtlich $\mathfrak{F}$-frei sind. Daher wird nach Satz 5

$$\dim K(\mathfrak{F}; g) \geq \dim K = \frac{\log 2}{\log (g^{q+r})} > 0.$$

b) Angenommen, es gibt keine zwei Sequenzen S und T mit den geforderten Eigenschaften. Da es in jeder g-adischen Ziffernfolge eine Sequenz der Länge i gibt, die unendlich oft auftritt, gibt es für jede Zahl $\varrho \in K(\mathfrak{F}; g)$ eine $\mathfrak{F}$-freie Sequenz J mit $|J| = i$, derart, daß ihre g-adische Ziffernfolge von der Form

$$U_0 J U_1 J U_2 J \ldots \tag{36}$$

ist, wobei die U_ν irgendwelche (möglicherweise leeren) $\mathfrak{F}$-freien Sequenzen sind. Dann ist $\mathfrak{Z}(JU_\nu)$ nach Hilfssatz 3 für $\nu = 1, 2, \ldots$ ein $\mathfrak{F}$-freier Zyklus, da wegen $\varrho \in K(\mathfrak{F}; g)$ für positives ν sowohl JU_ν als auch $U_\nu J$ eine $\mathfrak{F}$-freie Sequenz ist und $|J| = i$ gilt. Somit können nach Voraussetzung unter den Zyklen $\mathfrak{Z}(JU_\nu)$ keine zwei wesentlich verschiedenen sein. Setzt man nun bei der Darstellung (36) noch voraus, daß

$$t(U_\nu, J) = 0 \qquad (\nu = 0, 1, 2, \ldots)$$

ist, was offenbar keine Beschränkung der Allgemeinheit bedeutet, so ergibt sich, daß die Sequenzen JU_ν mit positivem ν sogar alle gleich sein müssen. Somit ist $U_1 = U_2 = \cdots = U$, und (36) hat die Form $U_0 J U J U J U \ldots$, d.h. die g-adische Entwicklung von ϱ ist von einer Stelle an periodisch. Daher ist $K(\mathfrak{F}; g)$ eine Menge von rationalen Zahlen und als solche höchstens abzählbar.

§ 5. Der Fall $m = 1$

12. Besteht die Menge $\mathfrak{F}$ nur aus einer einzigen Sequenz $F = f_1 \ldots f_i$, so schreiben wir $K(F; g) = K(f_1 \ldots f_i; g)$ statt $K(\mathfrak{F}; g)$. Für solche Mengen $K(F; g)$ lassen sich die Ergebnisse von § 4 noch verschärfen, wie im folgenden gezeigt werden soll.

Sei $\Xi = \Xi(F)$ der E-Fächer, dem die Menge $K(F; g)$ im Sinne von § 3 g-adisch zugeordnet ist. Dann gilt für die Klasse k von Ξ nach (21) die

Bedingung $k \leq i$; doch ist sogar $k = i$; denn sind E_1 und E_2 Sequenzen mit $t(E_1, F) \neq t(E_2, F)$, so kann man stets eine Sequenz G so angeben, daß von den Sequenzen E_1G und E_2G die eine F-frei, die andere dagegen F-haltig ist, und somit sind die zugehörigen Unterfächer verschieden, d.h. die Anzahl der verschiedenen Unterfächer von Ξ ist gleich der Anzahl der Werte von $t(E, F)$ für irgendwelche F-freien E. Daher numeriert man zweckmäßig die Unterfächer $\Xi_0, \ldots, \Xi_{i-1}$ so, daß für jeden Knotenpunkt A der Typ des Unterfächers $\Xi(A)$ mit dem Typ der zu A gehörenden Sequenz in bezug auf F übereinstimmt.

Die Matrix $\mathfrak{A}(\Xi(F))$, die dem E-Fächer Ξ im Sinne von § 2 zugeordnet werden kann, ist somit i-reihig und läßt sich leicht explizit angeben: das Matrixelement $a_{\varkappa\lambda}$ ist gleich der Anzahl der g-adischen Ziffern f mit

$$t(f_0 \ldots f_\lambda f, F) = \varkappa \qquad (\varkappa, \lambda = 0, \ldots, i-1;\ f_0 = \text{leere Sequenz}). \tag{37}$$

Daher ist $0 \leq a_{\varkappa\lambda} < g$ für alle $\varkappa$ und λ, ferner

$$\sum_{\varkappa=0}^{i-1} a_{\varkappa\lambda} = \begin{cases} g-1 & \text{für } \lambda = i-1 \\ g & \text{für } 0 \leq \lambda < i-1, \end{cases}$$

da bei festem λ von den g Sequenzen der Form $f_0 \ldots f_\lambda f$ entweder alle oder alle bis auf eine F-frei sind, je nachdem ob $\lambda = i-1$ ist oder nicht[15]).

Da der Grad eines Knotenpunktes vom Typ λ offenbar den Wert $1 + \sum_{\varkappa=0}^{i-1} a_{\varkappa\lambda}$ hat, ist somit jeder Knotenpunkt von Ξ vom Grad g oder $g+1$. Daher ist $g = 2$, falls in dem Fächer $\Xi(F)$ eine Rute[16]), d.h. ein unendlich langer Weg mit lauter Knotenpunkten vom Grad 2, auftritt. Ist dies bei einer dyadischen Sequenz $F = f_1 \ldots f_i$ der Fall, so nennen wir sie *direkt entartet* und die Sequenz $\tilde{F} = f_i \ldots f_1$ *indirekt entartet*. Als *entartet* bezeichnen wir eine dyadische Sequenz, wenn sie entweder direkt oder indirekt entartet ist.

Hilfssatz 4. a) Ist $g = 2$ und $|F| = 1$, so ist F entartet.

b) Ist $g = 2$ und $|F| = i > 1$, so ist F dann und nur dann direkt entartet, wenn

$$f_1 = f_2 = \cdots = f_{i-1} \neq f_i \tag{38}$$

gilt.

Beweis. Teil a) ist evident. Zum Beweis von Teil b) beachte man, daß, wenn F direkt entartet ist, alle Knotenpunkte auf der voraussetzungsgemäß in Ξ vorhandenen Rute vom Typ $i-1$ sein müssen. Demnach gilt, wenn f

[15]) Beispiel: Ist $g = 2$, $i = 6$, $F = 110111$, so erhält man auf Grund von (37)

$$\mathfrak{A}(\Xi(F)) = \begin{pmatrix} 1&1&0&0&0&0 \\ 1&0&1&0&0&0 \\ 0&0&1&1&0&0 \\ 1&0&0&0&1&0 \\ 1&0&0&0&0&1 \\ 0&0&0&1&0&0 \end{pmatrix}.$$

[16]) Dieser Begriff stammt von TH. KALUZA [7].

die von f_i verschiedene dyadische Ziffer ist, nach (37)

$$t(f_1 \dots f_{i-1} f, F) = i - 1,$$

woraus $f_1 \dots f_{i-1} = f_2 \dots f_{i-1} f$ und damit (38) folgt.

Eine direkt entartete dyadische Sequenz hat also die Form $11 \dots 10$ oder $00 \dots 01$, eine indirekt entartete die Form $01 \dots 11$ oder $10 \dots 00$. Die beiden dyadischen Sequenzen $F=0$ und $F=1$ der Länge 1 sind somit sowohl direkt als auch indirekt entartet. Wir lassen sie im folgenden außer Betracht, da im Sinne von (1) die Menge $K(1;2)$ leer ist und $K(0;2)$ nur das Element 1 hat. Im übrigen gilt der

Satz 8. Ist $i>1$ und $F=f_1 \dots f_i$ eine direkt entartete dyadische Sequenz, so ist die Menge

$$D = K(f_1 \dots f_i; g) - K(f_1 \dots f_{i-1}; g)$$

höchstens abzählbar.

Beweis. Ist $\varrho \in D$, so tritt bei ϱ die Sequenz $f_1 \dots f_{i-1}$, nicht aber die Sequenz $f_1 \dots f_i$ auf. Also müssen wegen (38) alle auf das erste solche Exemplar von $f_1 \dots f_{i-1}$ folgenden Ziffern von ϱ gleich der von f_i verschiedenen dyadischen Ziffer sein, d.h. ϱ ist von der Form

$$\varrho = 0, E f_1 \dots f_{i-1} f f f \dots, \tag{39}$$

wo E eine $(f_1 \dots f_{i-1})$-freie Sequenz ist, und das trifft für höchstens abzählbar viele ϱ zu. Wie man an Hand von (39) leicht bestätigt, ist D effektiv eine abzählbare Menge, wenn F von 01 verschieden ist; sonst ist D leer.

Setzt man (im Sinne von § 4) $\mathfrak{F} = \{F_1\} = \{F\}$ und $\mathfrak{P}_{11} = \mathfrak{P}(F) = \{p_1, \dots, p_c\}$, so gilt der

Satz 9. Sind $\gamma_1, \dots, \gamma_i$ die Lösungen der Gleichung[17])

$$z^i - g z^{i-1} + \sum_{\lambda=1}^{c} z^{p_\lambda} - g \sum_{\lambda=1}^{c} z^{p_\lambda - 1} + 1 = 0, \tag{40}$$

so gibt es unter ihnen eine einfache, etwa $\gamma_1 = \gamma(F)$, mit $\gamma(F) > 0$ derart, daß, wenn $\alpha = \dfrac{\log \gamma(F)}{\log g}$ und

$$\tau(F) = \prod_{j=2}^{i} \frac{g - \gamma_i}{\gamma - \gamma_j} \tag{41}$$

gesetzt wird, im Fall[18]) $\alpha > 0$

$$\frac{\gamma(F)}{g} \leq \{K(F;g)\}^{\alpha} \leq \tau(F) \tag{42}$$

und im Fall $\alpha = 0$ die Menge $K(F;g)$ höchstens abzählbar, also in beiden Fällen $\dim K(F;g) = \alpha$ ist.

13. Für den Beweis benötigen wir die folgenden Hilfssätze:

Hilfssatz 5. Die charakteristische Gleichung der Matrix $\mathfrak{A}(\Xi(F))$ ist (40).

[17]) Bei leerem $\mathfrak{P}$ sind die Summen durch 0 zu ersetzen.

[18]) $\{M\}^{\alpha}$ bezeichnet das α-dimensionale Hausdorffsche Maß der Menge M; vgl. Nr. 5.

Beweis. Sei $x(n)$ die Anzahl der F-freien Sequenzen E mit $|E|=n$ und $t(E,F)=i-1$. Dann gilt für die Anzahl $h(n)$ der F-freien Sequenzen der Länge n nach (24)

$$h(n+1)=g\,h(n)-x(n) \qquad (n\geq i-1). \tag{43}$$

Ist $\mathfrak{P}(F)$ leer, so ist nach (26)

$$h(n+1)=g\,h(n)-h(n-i+1) \qquad (n\geq i-1). \tag{44}$$

Andernfalls folgt aus (30)

$$x(n)=h(n-i+1)-\sum_{\lambda=1}^{c}x(n-i+p_\lambda).$$

Setzt man hierin gemäß (43)

$$x(n-i+p_\lambda)=g\,h(n-i+p_\lambda)-h(n+1-i+p_\lambda),$$

so erhält man die für $n\geq i-1$ gültige Gleichung

$$h(n+1)=g\,h(n)-\sum_{\lambda=1}^{c}h(n+1-i+p_\lambda)+g\sum_{\lambda=1}^{c}h(n-i+p_\lambda)-h(n-i+1). \tag{45}$$

Aus (44) und (45) folgt die Behauptung nach Satz 1.

Hilfssatz 6. Ist $F=f_1\dots f_i$ nicht entartet, so gibt es zu jedem F-freien E mit $t(E,F)=i-1$ ein E' mit $|E'|=i-1$, so daß EE' eine F-freie Sequenz mit $t(EE',F)=0$ wird.

Beweis. Für $i=1$ ist die Behauptung bei jedem F-freien E' erfüllt. Ist $i>1$, so genügt es offenbar, den Beweis für die Sequenz $E=f_1\dots f_{i-1}$ zu erbringen. Falls $g=2$ und $f_1=f_i$ oder falls $g\geq 3$ ist, gibt es eine Ziffer e mit $f_1\neq e\neq f_i$. Daher hat die Sequenz $E'=ee\dots e$ mit $|E'|=i-1$ die geforderten Eigenschaften.

In dem übrigbleibenden Fall $g=2$, $f_1\neq f_i$ setzen wir voraus, daß

$$f_1=0, \quad f_i=1 \tag{46}$$

ist; der umgekehrte Fall ergibt sich aus den folgenden Betrachtungen, wenn überall 0 und 1 vertauscht werden. Da für $i=2$ und $i=3$ alle dyadischen Sequenzen mit (46), nämlich 01 sowie 001 und 011, entartet sind, kann ferner $i\geq 4$ vorausgesetzt werden. Wir bezeichnen mit $a\geq 1$ die Länge der längsten „Kette" von aufeinanderfolgenden Einsen in F und mit $b\geq 0$ den Typ der Sequenz E in bezug auf die Sequenz $N=00\dots00$ mit $|N|=i-1$. Dann ist, da F nicht entartet ist, $a<i-1$ und $b<i-1$, wie aus (46) folgt. Wir unterscheiden nun zwei Fälle:

a) Enthält F nur eine einzige Kette von Einsen, so muß F wegen (46) von der Form

$$f_j=\begin{cases}0 & \text{für } 1\leq j\leq i-a\\ 1 & \text{für } i-a<j\leq i\end{cases}$$

sein. Hierin ist $a>1$ und $i-a>1$, weil F sonst entartet wäre. Setzt man daher $E'=01$, so ist $EE'=f_1\dots f_{i-1}01=\underbrace{00\dots0}_{i-a}\underbrace{11\dots1}_{a-1}01$ und somit sicher $t(EE',F)=0$.

b) Enthält F mehr als eine Kette von Einsen, so ist $a\leq i-3$; denn außer der Kette der Länge a gibt es unter den i Ziffern von F wenigstens noch: eine weitere Kette von mindestens einer Eins, mindestens eine Null zwischen beiden Ketten und die Ziffer $f_1=0$. Folglich wird, wenn wir $E'=0\underbrace{11\dots1}_{a+1}$ setzen, $|E'|=a+2\leq i-1$. Dann ist EE', wie gezeigt werden soll, F-frei. Ein etwa auftretendes Exemplar von F könnte nicht aus den letzten i Ziffern von EE' bestehen, weil F keine Kette von Einsen der Länge $a+1$ enthält. Also müßten auch Einsen dieses Exemplars von F in E liegen. Das ist jedoch ausgeschlossen, weil die am weitesten rechts gelegene Kette von Nullen in EE' nach der Definition von b die Länge $b+1$ hat, wegen ihrer Lage in F aber die Länge b haben müßte.

Wäre nun $t(EE',F)>0$, so müßte wegen $f_1=0$ offenbar $t(EE',F)>a+1$ sein, was aber unmöglich ist, da F sonst eine Kette von $a+1$ Einsen enthalten müßte. Also ist, wie gefordert, $t(EE',F)=0$.

Hilfssatz 7. Ist $F=f_1\dots f_i$ nicht entartet und $n\geq 3(i-1)$, so gibt es zu jeder F-freien Sequenz E und jedem $\varkappa$ mit $0\leq\varkappa\leq i-1$ ein E' mit $|E'|=n$, so daß EE' ebenfalls F-frei und $t(EE',F)=\varkappa$ ist.

Beweis. Ist $t(E,F)=\lambda$, so setzen wir $E_1=f_{\lambda+1}\dots f_{i-1}$, so daß also $|E_1|\leq i-1$ und $t(EE_1,F)=i-1$ ist. Folglich gibt es nach Hilfssatz 6 ein E_2 mit $|E_2|\leq i-1$, so daß $t(EE_1E_2,F)=0$ ist. Bezeichnen wir, wenn f eine Ziffer $\neq f_1$ ist, mit E_3 die Sequenz $00\dots0$ der Länge $n-\varkappa-|E_1|-|E_2|$, die ja wegen $n\geq 3(i-1)$ nichtnegativ ausfällt, so ist $t(EE_1E_2E_3,F)=0$ und ferner $|E_1E_2E_3|=n-\varkappa$. Daher hat, wenn wir $E_4=f_1\dots f_\varkappa$ setzen, die Sequenz $E'=E_1E_2E_3E_4$ die geforderten Eigenschaften.

Hilfssatz 8. Ist $n\geq 3(i-1)$, so sind alle Elemente der Matrix $\mathfrak{A}(\Xi(F))^n$ positiv.

Beweis. Bezeichnet man die Elemente von $\mathfrak{A}(\Xi(F))^n$ wieder mit $a_{\varkappa\lambda}^{(n)}$ $(\varkappa,\lambda=0,\dots,i-1)$, so ergibt sich aus (2) auf Grund der Konstruktion des Fächers $\Xi(F)$, daß für jedes feste F-freie E mit $t(E,F)=\lambda$ das Matrixelement $a_{\varkappa\lambda}^{(n)}$ gleich der Anzahl der F-freien Sequenzen der Form EE' mit $|E'|=n$ und $t(EE',F)=n$ ist. Nach Hilfssatz 7 ist diese Anzahl für $n\geq 3(i-1)$ positiv.

Hilfssatz 9. Ist $\mathfrak{M}$ eine Matrix mit positiven Elementen, so gibt es unter den Eigenwerten von $\mathfrak{M}$ einen, der positiv-reell, einfach und von größerem Absolutbetrag als alle übrigen ist.

Beweis. Siehe Frobenius [4].

Hilfssatz 10. Die Eigenwerte der Matrix $\mathfrak{M}^n$ sind die n-ten Potenzen der Eigenwerte der Matrix $\mathfrak{M}$.

Beweis. Siehe z.B. MacDuffee [2], Theorem 16.1.

Hilfssatz 11. In der Bezeichnungsweise von Satz 9 ist, falls $\gamma(F) \neq 1$ ist,

$$h(n) = \gamma(F)^n \left(\tau(F) + o(1)\right).$$

Beweis. a) Sei F nicht entartet. Dann ist nach den Hilfssätzen 5, 8, 9 und 10 für $n \geq 3(i-1)$, wenn $\xi_1, \ldots, \xi_q$ die verschiedenen Eigenwerte von $\mathfrak{A}(\Xi(F))^n$ und $v_1, \ldots, v_q$ ihre Vielfachheiten $\left(\sum_{j=1}^{q} v_j = i\right)$, ferner $\gamma_1, \ldots, \gamma_{q'}$ die verschiedenen Eigenwerte von $\mathfrak{A}(\Xi(F))$ und $v_1', \ldots, v_{q'}'$ deren Vielfachheiten sind, bei geeigneter Numerierung

$$q = q', \; v_j = v_j' \quad \text{und} \quad \xi_j = \gamma_j^n \qquad (j = 1, \ldots, q),$$

ferner ist ξ_1 positiv-reell, und es gilt $v_1 = 1$ sowie

$$|\xi_j| < \xi_1 \qquad (j = 2, \ldots, q).$$

Daraus folgt, daß auch $\gamma_1 = \gamma(F)$ positiv-reell ist und

$$|\gamma_j| < \gamma_1 \qquad (j = 2, \ldots, q) \tag{47}$$

gilt. Wie aus (22) folgt, läßt sich $h(n)$ somit in der Form

$$h(n) = \sigma \gamma_1^n + P_2(n)\gamma_2^n + \cdots + P_q(n)\gamma_q^n \qquad (n = 0, 1, \ldots) \tag{48}$$

darstellen, wo $P_j(n)$ ein Polynom von einem Grade $< v_j$ ist. Aus (48) und (47) folgt

$$h(n) = \gamma_1^n \left(\sigma + P_2(n)\left(\frac{\gamma_2}{\gamma_1}\right)^n + \cdots + P_q\left(\frac{\gamma_q}{\gamma_1}\right)^n\right) = \gamma_1^n \left(\sigma + o(1)\right). \tag{49}$$

Zur Bestimmung von σ setzen wir zunächst voraus, daß die Vielfachheiten v_j alle den Wert 1 haben, also $q = i$ ist und somit die Polynome $P_j(n)$ sämtlich konstant sind. Da für $0 \leq n \leq i-1$ offenbar $h(j) = g^j$ ist, wird also $\sigma = \frac{\Delta_1}{\Delta}$, wenn wir

$$\Delta = \begin{vmatrix} g^0 & \gamma_2^0 & \cdots & \gamma_i^0 \\ g^1 & \gamma_2^1 & \cdots & \gamma_i^1 \\ \vdots & & & \vdots \\ g^{i-1} & \gamma_2^{i-1} & \cdots & \gamma_i^{i-1} \end{vmatrix} \quad \text{und} \quad \Delta_1 = \begin{vmatrix} \gamma_1^0 & \gamma_2^0 & \cdots & \gamma_i^0 \\ \gamma_1^1 & \gamma_2^1 & \cdots & \gamma_i^1 \\ \vdots & & & \vdots \\ \gamma_1^{i-1} & \gamma_2^{i-1} & \cdots & \gamma_i^{i-1} \end{vmatrix}$$

setzen; denn Δ ist als Differenzenprodukt der γ_j in diesem Fall sicher $\neq 0$. Ferner ist $\Delta_1 = \prod_{i=2}^{i} (g - \gamma_j) \prod_{j,l=2}^{i} (\gamma_j - \gamma_l)$ und daher

$$\sigma = \prod_{j=2}^{i} \frac{g - \gamma_j}{\gamma_1 - \gamma_j} = \tau(F).$$

Wie eine Stetigkeitsbetrachtung zeigt, bleibt diese Gleichung auch bestehen, wenn die Wurzeln $\gamma_2, \ldots, \gamma_i$ nicht alle verschieden sind, und somit folgt (46) aus (49). Insbesondere ist $\tau(F) > 0$, da im Fall $\tau(F) = 0$ eine der Wurzeln $\gamma_i = g$ sein müßte, während doch die linke Seite von (40) für $z = g$ den Wert 1 annimmt, und da der Fall $\tau(F) < 0$ wegen $h(n) \to +\infty$ ausgeschlossen ist.

b) Ist F direkt entartet, so ist $\mathfrak{P}(F)$ leer und daher, wenn man die linke Seite von (40) mit $\varphi(F;z)$ bezeichnet,

$$\varphi(f_1 \dots f_i; z) = z^i - 2z^{i-1} + 1 = \left(z^{i-1} - \sum_{j=0}^{i-2} z^j\right)(z-1) = \varphi(f_1 \dots f_{i-1}; z)\,(z-1).$$

Daher ist nach (48)

$$h(f_1 \dots f_i; n) = h(f_1 \dots f_{i-1}; n) + c_6.$$

Hieraus folgt im Fall $i \geq 3$, wo die Sequenz $f_1 \dots f_{i-1}$ nicht entartet und also schon unter a) mitbehandelt ist, die Behauptung. Die Fälle $i=2$ und $i=1$ führen auf den oben ausgeschlossenen Wert $\gamma_1 = 1$.

c) Ist F indirekt entartet, also $\tilde{F}$ direkt entartet, so folgt die Behauptung aus b), da die Funktionen $h(F;n)$ und $h(\tilde{F};n)$ derselben Differenzengleichung mit denselben Anfangsbedingungen genügen und folglich identisch sind.

Hilfssatz 12. Für die Funktionen $h_\varkappa(F;n)$ des Fächers $\Xi(F)$ im Sinne von Satz 1 ist

$$h_\varkappa(F;n) \leq h(F;n) \qquad (\varkappa = 0, \dots, i-1;\ h_0(F;n) = h(F;n)),$$

und es gibt eine Konstante $\tau_\varkappa$ mit $0 \leq \tau_\varkappa \leq \tau(F)$ und $h_\varkappa(n) = \gamma_1^n\,(\tau_\varkappa + o(1))$.

Beweis. Sei E mit $t(E,F) = \varkappa$ eine feste F-freie Sequenz. Dann ist $h_\varkappa(F;n)$ die Anzahl der Sequenzen E' mit $|E'| = n$ und F-freiem EE'. Da jedes solche E' selbst F-frei ist, folgt $h_\varkappa(F;n) \leq h(F;n)$. Die zweite Behauptung läßt sich unter Verwendung von Satz 1 genau so wie Hilfssatz 11 beweisen.

Hilfssatz 13[19]**.** Ist $0 < \delta \leq 1$ und M eine abgeschlossene beschränkte lineare Punktmenge mit $\{M\}^\delta < \infty$, so ist dann und nur dann $\{M\}^\delta > 0$, wenn es eine nichtnegative additive Funktion $\Phi(R)$ von halboffenen Intervallmengen[20]) gibt, für die folgendes gilt:

a) Für jede halboffene Intervallmenge R ist $\Phi(R) \leq c_7 |R|^\delta$.

b) Für jede halboffene Intervallmenge R mit $R' \supseteq M$ ist $\Phi(R) \geq c_8$.

Dann ist nämlich $\{M\}^\delta \geq \frac{c_8}{c_7}$.

Beweis. Siehe Moran [8], Theorem I.

14. Nach diesen Hilfsbetrachtungen erbringen wir nunmehr den

Beweis von Satz 9. Ist $\gamma(F) = 1$, so folgt aus (40), daß

$$1 - g + c(1-g) + 1 = 0,$$

d.h. $(g-1)(c+1) = 1$ und folglich $g = 2$, $c = 0$ ist. Also ist $\mathfrak{P}(F)$ leer, und (40) nimmt die Gestalt $z^i - 2z^{i-1} + 1 = 0$ an. Dieses Polynom hat offenbar genau dann $\gamma_1 = 1$ als größte Nullstelle, wenn $i = 1$ oder $i = 2$ ist. Man erhält so die vier Mengen[21])

(50) $$K(1;2),\ K(0;2),\ K(01;2) \text{ und } K(10;2),$$

[19]) Wir spezialisieren den Satz hier auf den Fall α-dimensionaler Maße und linearer Mengen.

[20]) Unter einer halboffenen Intervallmenge wird eine Menge von endlich vielen halboffenen Intervallen verstanden.

[21]) Vgl. [12], Satz 5.

von denen die erste leer und die vierte abzählbar sind, während die zweite und dritte nur das Element 1 enthalten.

Sei im folgenden $\gamma(F)=\gamma_1>1$ vorausgesetzt[22]). Dann ist, da die in Nr. 7 eingeführte Menge $I_n(F)'$ eine g^{-n}-Überdeckung von $K(F;g)$ bildet, nach Hilfssatz 11

$$(51)\qquad \{K(F;g)\}^\alpha \leq \lim_{n\to\infty} h(n)\, g^{-n\alpha} = \lim_{n\to\infty} \gamma_1^{-n}\, g^{-n\alpha}\big(\tau + o(1)\big) = \tau(F).$$

Zu dem noch fehlenden Nachweis[23]) von

$$(52)\qquad \{K(F;g)\}^\alpha \geq \frac{\gamma(F)}{g}$$

soll gezeigt werden, daß Hilfssatz 13 auf die abgeschlossene Hülle $K(F;g)^*$ von $K(F;g)$ mit $\delta=\alpha$ anwendbar ist. Da die Voraussetzung $\{M\}^\delta<\infty$ nach (51) erfüllt ist, bleibt nur die Existenz einer Funktion $\Phi(R)$ mit den geforderten Eigenschaften zu zeigen. Dazu definieren wir für jedes Intervall $\mathfrak{i}_n^\nu$ die Funktion $\psi(\mathfrak{i}_n^\nu)=\gamma_1^{-n}$ und setzen für jede halboffene Intervallmenge R

$$\Phi(R) = \lim_{n\to\infty} \sum_\nu \psi(\mathfrak{i}_n^\nu),$$

wobei über alle $\mathfrak{i}_n^\nu \in I_n$ summiert wird, deren innere Punkte sämtlich in R liegen. Um zu zeigen, daß dieser Grenzwert existiert, gehen wir in mehreren Schritten vor:

1) Sei R ein Intervall von I_j. Ist R dann etwa vom Typ $\varkappa$, so ist die Anzahl der darin enthaltenen Intervalle von I_n $(n\geq j)$ offenbar $h_\varkappa(n-j)$, und folglich existiert nach Hilfssatz 12 der Grenzwert

$$\Phi(R) = \lim_{n\to\infty} h_\varkappa(n-j)\,\gamma_1^{-n} = \tau_\varkappa \gamma_1^{-j} \leq \tau(F)\,\gamma_1^{-j}.$$

2) Sei R eine endliche Untermenge von $I=\bigvee_{\nu=1}^{\infty} I_n(F)$, also etwa $R=\{\mathfrak{i}_{j_1}^{m_1},\ldots,\mathfrak{i}_{j_w}^{m_w}\}$. Dann existiert $\Phi(R)$, und es ist

$$\Phi(R) = \sum_{\nu=1}^{w} \Phi(\mathfrak{i}_{j_\nu}^{m_\nu}),$$

wie sich auf Grund von 1) ergibt.

3) R bestehe aus einem beliebigen Intervall. Dann wählen wir j so groß, daß $g^{-j}<|R'|$ wird. Für $n\geq j$ ist wegen 2) und 1), da es höchstens zwei Intervalle von $I_j(F)$ gibt, die mit R' einen nichtleeren Durchschnitt haben, ohne ganz in R' enthalten zu sein, nach Hilfssatz 12

$$0 \leq \sum_{\mathfrak{i}_n\in I_n\wedge R'} \Phi(\mathfrak{i}_n) - \sum_{\mathfrak{i}_j\in I_j\wedge R'} \Phi(\mathfrak{i}_j) \leq 2\Phi(\mathfrak{i}_j) \leq 2\tau(F)\,\gamma_1^{-j},$$

und somit existiert der Grenzwert $\Phi(R)$, da $\lim_{j\to\infty} 2\tau(F)\,\gamma_1^{-j}=0$ ist.

4. Sei R irgendeine halboffene Intervallmenge, etwa $R=\{\mathfrak{r}_1,\ldots,\mathfrak{r}_l\}$, wobei diese $\mathfrak{r}_\nu$ halboffene, sich nicht überschneidende Intervalle sind. Dann

[22]) Der Fall $\gamma_1<1$ kann nicht eintreten, da nach (40) stets $\gamma(F)\geq g-1$ ist.

[23]) Die folgenden Betrachtungen bilden eine Verbesserung von [12], Nr. 4.

kann man ohne Beschränkung der Allgemeinheit annehmen, daß zwischen je zweien dieser Intervalle jeweils ein nicht in R enthaltenes Intervall liegt, und man kann somit die Überlegung 3) auf jedes $\mathfrak{r}_\nu$ gesondert anwenden. Daraus folgt, daß $\Phi(R)$ existiert und $\Phi(R) = \sum_{\nu=1}^{t} \Phi(\mathfrak{r}_\nu)$ ist.

Damit ist die Funktion $\Phi(R)$ für alle R definiert, und eine einfache Betrachtung zeigt, daß sie auch additiv ist. Wir wollen nun zeigen, daß die Voraussetzungen a) und b) von Hilfssatz 13 mit $c_7 = \tau \frac{g}{\gamma_1}$, $c_8 = \tau$ erfüllt sind.

Zu a). Sei für die halboffene Intervallmenge R die Zahl j so bestimmt, daß

$$g^{-(j+1)} \leq |R'| < g^{-j} \tag{53}$$

ist. Da R' von einem Intervall $\mathfrak{i}$ der Länge $|R'|$ überdeckt wird, ist, wenn etwa $\mathfrak{i}_{j+1}^{m_1}, \ldots, \mathfrak{i}_{j+1}^{m_l}$ die in $\mathfrak{i}$ enthaltenen Intervalle von I_{j+1} sind, nach 1)

$$\Phi(R) = \sum_{\nu=1}^{l} \Phi(\mathfrak{i}_{j+1}^{m_\nu}) \leq l\, \tau(F)\, \gamma_1^{-(j+1)}.$$

Somit wird, da trivialerweise $l \leq \frac{|R'|}{g^{-(j+1)}}$ ist, wegen $\gamma(F) = g^\alpha$

$$\Phi(R) \leq \tau\, g^{1-\alpha} |R'|^\alpha = \tau \frac{g}{\gamma_1} |R'|^\alpha,$$

woraus wegen (53)

$$\Phi(R) \leq \tau\, g^{1-\alpha} |R'|^\alpha = \tau \frac{g}{\gamma_1} |R'|^\alpha$$

folgt.

Zu b). Sei R die Menge der halboffenen Intervalle $\mathfrak{r}_1, \ldots, \mathfrak{r}_t$, und sei $R' \supseteq K(F; g)^*$. Dann gibt es bei genügend großem n höchstens $2t$ Intervalle von I_n, die nicht vollständig in R' enthalten sind. Daher wird nach Hilfssatz 11

$$\Phi(R) = \lim_{n \to \infty} \frac{h(n) - 2t}{g^{n\alpha}} = \tau(F) > 0.$$

Folglich ergibt sich nach Hilfssatz 13

$$\{K(F; g)\}^\alpha = \{K(F; g)^*\}^\alpha \geq \frac{c_8}{c_7} = \frac{\gamma_1}{g},$$

womit (52) bewiesen ist.

§ 6. Analoge Sätze für Mengen natürlicher Zahlen

15. Die Untersuchungen der beiden vorangehenden Paragraphen lassen sich in naheliegender Weise auf Mengen $\mathfrak{M}$ natürlicher Zahlen übertragen, wobei sich statt der Aussagen über die Dimension von Punktmengen solche über das asymptotische Verhalten der Anzahlfunktionen $M(x)$ dieser Mengen $\mathfrak{M}$, d.h. der Anzahl der Elemente $m \in \mathfrak{M}$ mit $m \leq x$, ergeben. Es gilt der

Satz 10. Sei Ξ ein E-Fächer, dessen Funktion $h(n)$ der Ungleichung (7) genügt, und sei $M = \bigwedge_{n=1}^{\infty} I'_n$ eine ihm g-adisch zugeordnete Punktmenge. Wird dann jeder der Sequenzen $E_n^r = e_1 \ldots e_n$ die ganze Zahl $e_{n\,r} = \sum_{j=1}^{n} e_j\, g^{n-j}$

zugeordnet und die Menge aller dieser $e_{n,r}$ mit $\mathfrak{M}$ bezeichnet, so ist

$$M(x) = x^{\dim M + o(1)}.$$

Beweis. Für jede Sequenz E_n^r, die zu einem Intervall $i_n^r \in I_n$ gehört, ist offenbar $0 \leq e_{n,r} < g^n$. Da auch umgekehrt jeder Zahl $m \in \mathfrak{M}$ mit $0 \leq m < g^n$, wenn sie gegebenenfalls durch Voranstellen von Nullen auf n Ziffern verlängert wird, eine Sequenz E mit $|E| = n$ entspricht, die zu einem Intervall $i_n^r \in I_n$ gehört, ist

$$h(n) = M(g^n - 1) \qquad (n = 0, 1, \ldots).$$

Folglich gilt für jede ganze Zahl x mit $g^n \leq x < g^{n+1}$, also mit $n = \left[\frac{\log x}{\log g}\right]$, die Beziehung

$$h(n) \leq M(x-1) < h(n+1). \tag{54}$$

Daraus ergibt sich nach Satz 2, da ja in der dortigen Bezeichnungsweise

$$h(n) = e^{n(\chi + o(1))} = g^{n(\dim M + o(1))}$$

ist, wie behauptet,

$$M(x) = g^{\frac{\log x}{\log g}(\dim M + o(1))} = x^{\dim M + o(1)}.$$

Ist $\mathfrak{F} = \{F_1, \ldots, F_m\}$ eine Menge g-adischer Sequenzen im Sinne von Nr. 7, so gilt für die Menge $\mathfrak{N}(\mathfrak{F}; g)$ der nichtnegativen ganzen Zahlen mit $\mathfrak{F}$-freier g-adischer Entwicklung der

Satz 11. Die Anzahlfunktion $K(\mathfrak{F}; g; x) = K(x)$ genügt der asymptotischen Gleichung

$$K(x) = x^{\frac{\log \gamma(\mathfrak{F})}{\log g} + o(1)}$$

Beweis. Sind die ersten Ziffern $f_1^1, f_1^2, \ldots, f_1^m$ sämtlich von Null verschieden, so entsteht aus jeder Zahl $k \in \mathfrak{N}(\mathfrak{F}; g)$ mit $k < g^n$, wenn man ihre g-adische Entwicklung durch Voranstellen von Nullen auf n Ziffern verlängert, eine $\mathfrak{F}$-freie Sequenz der Länge n, und daher ist

$$K(g^n - 1) = h(n) \qquad (n = 0, 1, \ldots). \tag{55}$$

Gibt es dagegen eine Sequenz F_μ mit $f_1^\mu = 0$, ist also

$$\max_{\mu = 1, \ldots, m} t(\underbrace{00 \ldots 0}_{i-1}, F_\mu) = \lambda > 0,$$

so geht die $\mathfrak{F}$-Freiheit von k durch diesen Prozeß verloren, falls die ersten $i - \lambda$ Ziffern von k eine der Sequenzen $f_\mu^{\lambda+1} f_\mu^{\lambda+2} \ldots f_\mu^i$ bilden, bei denen $f_\mu^1 \ldots f_\mu^\lambda = 0 \ldots 0$ ist. Bezeichnet man dann mit $A(s)$ die Anzahl der $\mathfrak{F}$-freien nichtnegativen Zahlen $k < g^n$ mit s Ziffern, die nach Voranstellen von $n - s$ Nullen nicht mehr $\mathfrak{F}$-frei sind, so wird wegen $\lambda \leq i - 1$

$$A(s) \leq m\, h(s - i + \lambda) \leq m\, h(s-1)$$

und ferner

$$K(g^n - 1) - h(n) \leq \sum_{s=1}^{n-1} A(s) \leq m \sum_{s=0}^{n-1} h(s) \leq m n\, h(n).$$

So erhält man statt (55)

$$h(n) \leq K(g^n - 1) \leq (m+1)\, n\, h(n), \tag{56}$$

und daraus folgt die Behauptung, da nach Satz 3 sowohl

$$h(n) = g^{n(\dim K(\mathfrak{F};g) + o(1))} \quad \text{als auch} \quad (m+1)\, n\, h(n) = g^{n(\dim K(\mathfrak{F};g) + o(1))}$$

ist.

Folgerung. Sei F eine Sequenz mit[24]) $f_1 \neq 0$ und $\mathfrak{N}(F;g)$ die Menge der nichtnegativen ganzen Zahlen mit F-freier g-adischer Entwicklung. Dann gilt nach (54) für ihre Anzahlfunktion $K(x)$, wenn wieder $g^n \leq x < g^{n+1}$ angenommen wird,

$$\frac{h(n)}{(g^{n+1}-1)^\alpha} \leq \frac{K(x-1)}{(x-1)^\alpha} \leq \frac{h(n+1)}{(g^n-1)^\alpha},$$

und wegen $h(n) \simeq \tau g^{n\alpha} = \gamma^n$ (Hilfssatz 11) folgt

$$\lim_{x\to\infty} \frac{h(n)}{(g^{n+1}-1)^\alpha} = \frac{\tau}{g^\alpha} = \frac{\tau}{\gamma} \quad \text{und} \quad \lim_{x\to\infty} \frac{h(n+1)}{(g^n-1)^\alpha} = \tau\gamma,$$

so daß sich

$$\frac{\tau}{\gamma} \leq \underline{\lim_{x\to\infty}} \frac{K(x)}{x^\alpha} \leq \overline{\lim_{x\to\infty}} \frac{K(x)}{x^\alpha} \leq \tau\gamma \tag{57}$$

ergibt.

Literatur

[1] Best, E.: On sets of fractional dimension. III. Proc. London Math. Soc. 2, **47**, 436—454 (1942). — [2] MacDuffee, C. C.: The Theory of Matrices. Erg. Math., 2, 2. Aufl. New York 1946. — [3] Eggleston, H. G.: Sets of fractional dimensions which occur in some problems of number theory. Proc. London Math. Soc. 2, **54**, 42—93 (1951). — [4] Frobenius, G.: Über Matrizen aus positiven Elementen. Sitzgsber. preuß. Akad. Wiss., Physik.-math. Kl. **26**, 471—476 (1908). — [5] Good, I. J.: The fractional dimensional theory of continued fractions. Proc. Cambridge Phil. Soc. **37**, 199—228 (1941). — [6] Hardy, G., u. J. Littlewood: The fractional part of $n^k\Theta$. Acta Math. **37**, 155—191 (1914). — [7] Kaluza, Th.: Struktur- und Mächtigkeitsuntersuchungen an gewissen unendlichen Graphen mit einigen Anwendungen auf lineare Punktmengen. Math. Ann. **122**, 235—258 (1950). — [8] Moran, P. A. P.: Additive functions of intervals and Hausdorff measure. Proc. Cambridge Phil. Soc. **42**, 15—23 (1946). — [9] Nörlund, N. E.: Vorlesungen über Differenzenrechnung. Berlin 1924. — [10] Volkmann, B.: Über Klassen von Mengen natürlicher Zahlen. J. reine angew. Math. **190**, 199—230 (1952). — [11] Volkmann, B.: Über Hausdorffsche Dimensionen von Mengen, die durch Zifferneigenschaften charakterisiert sind. I. Math. Z. **58**, 284—287 (1953). — [12] Volkmann, B.: Über Hausdorffsche Dimensionen von Mengen, die durch Zifferneigenschaften charakterisiert sind. III. Math. Z. **59**, 259—270 (1953).

[24]) Dieser Zusatz ist auch bei [12], Nr. 6 zu machen.

Mainz, Mathematisches Institut der Universität und University of Utah, Department of Mathematics

(Eingegangen am 13. Februar 1956)

MORITA, K., and H. TACHIKAWA
Math. Zeitschr. Bd. 65, S. 414—428 (1956)

Character modules, submodules of a free module, and quasi-Frobenius rings

By
KIITI MORITA and HIROYUKI TACHIKAWA

INTRODUCTION. Let A be an associative ring with a unit element 1. By a left or right A-module we shall always mean one on which the unit element of A acts as the identity operator. In the present paper we shall investigate some properties of character modules, submodules and residue class modules of finite free A-modules, and give some characterizations of quasi-Frobenius rings.

Let U be a double A-module; that is, U is a left and right A-module, and $a(ub)=(au)b$ for $a, b\in A$ and $u\in U$. For a left A-module L, an A-left homomorphism of L into U will be called a U-character of L. The set of all U-characters of L forms a right A-module under a suitable definition. We shall call this module the U-character module[1]) of L. We may define the U-character module of a right A-module similarly.

Let the U-character modules of any simple left A-module and any simple right A-module be simple. Then our first theorem asserts that if a left A-module L and a right A-module R form an orthogonal pair with respect to U (for the definition, cf. § 1) and one of these modules has a composition series, then each of these modules is isomorphic to the U-character module of the other and the well known relation between the annihilators is valid. As an immediate consequence we have a theorem of M. HALL and the duality of character modules for quasi-Frobenius rings.

Now let A be a ring satisfying the minimum condition and possessing a unit element. In § 3 we shall prove the following theorem: If L and L' are two left A-submodules of a finite free left A-module $A^{(n)}$ with n generators such that $A^{(n)}/L$ is A-isomorphic to $A^{(n)}/L'$, then L and L' are A-isomorphic to each other. Special cases of this theorem have been obtained by GASCHÜTZ and KAWADA[2]).

The converse of this theorem does not hold in general, but we can establish the following theorem in § 4: In order that an A-isomorphism $L\cong L'$ for any two A-submodules of $A^{(n)}$ imply an A-isomorphism $A^{(n)}/L\cong A^{(n)}/L'$, $n=1, 2, \ldots$, it is necessary and sufficient that A be a quasi-Frobenius ring.

[1]) AZUMAYA [8] called it the conjugate module of L in case $U=A$. To emphasize the analogy to character groups of topological groups we adopt here this denomination (cf. [10] and § 1 below).

[2]) GASCHÜTZ [2] has proved this theorem for the case where A is a group ring over a prime field with characteristic $p\neq 0$, and KAWADA [4] has proved it for the case $n=1$.

Indeed, in case A is quasi-Frobenius, it is shown that every A-isomorphism $L \cong L'$ can be extended to an A-isomorphism of $A^{(\aleph)}$ onto itself.

The notions of (M_0)-modules and (M_u)-modules have been recently introduced by GASCHÜTZ [1] for the case of group rings, and NAGAO-NAKAYAMA [6] have determined the structure of these modules for the case of algebras. In § 5 we shall determine the structure of (M_u)-modules for quasi-Frobenius rings and furthermore show that a ring A is an (M_u)-module if and only if A is quasi-Frobenius[3]).

In case A is an algebra over a field K each of the theorems in § 3 and § 4 admits a representation-theoretical interpretation. We can give further a representation-theoretical characterization for quasi-Frobenius algebras: We shall prove in § 6 that every representation of A in K is embedded in a direct sum of a suitable finite number of left regular representations (i.e. every left representation module is A-isomorphic to a submodule of a suitable finite free left A-module), if and only if A is quasi-Frobenius.

§ 1. Character modules

Let A be a ring and U a double A-module. If L is a left A-module, the set of all U-characters of L forms a right A-module under the following definitions of addition and multiplication:

$$(\varphi + \psi)(x) = \varphi(x) + \psi(x), \quad \text{for} \quad x \in L,$$
$$(\varphi a)(x) = \varphi(x)\, a, \quad \text{for} \quad x \in L,\ a \in A.$$

This A-module is the U-character module of L and will be denoted by $\mathrm{Char}_U L$. The U-character module of a right A-module is similarly defined; it is a left A-module.

Let L and R be a left and a right A-module respectively. If, for any element x of L and any element y of R there corresponds an element u of U, denoted by (x, y), such that

$$(x + x', y) = (x, y) + (x', y)$$
$$(x, y + y') = (x, y) + (x, y')$$
$$(a x, y) = a(x, y), (x, y a) = (x, y)\, a,$$

for any $x, x' \in L$; $y, y' \in R$; $a \in A$, then we shall say that L and R form a pair with respect to U. In this case we put

$$\mathrm{Ann}(L_0; R) = \{y \mid (x, y) = 0, \quad \text{for all } x \in L_0\},$$
$$\mathrm{Ann}(R_0; L) = \{x \mid (x, y) = 0, \quad \text{for all } y \in R_0\},$$

for $L_0 \subseteq L$ and $R_0 \subseteq R$. If

$$\mathrm{Ann}(L; R) = 0, \quad \mathrm{Ann}(R; L) = 0,$$

[3]) This is a generalization of the result of NAGAO-NAKAYAMA obtained for the case of algebras.

we shall say that L and R form an orthogonal pair with respect to U. We shall prove the following theorem.

THEOREM 1.1. *Let the U-character modules of any simple left A-module and of any simple right A-module be simple. Suppose that a left A-module L and a right A-module R form an orthogonal pair with respect to U and one of L and R has a composition series. Then we have the following conclusions:*

(a) $\mathrm{Char}_U L \cong R, \quad \mathrm{Char}_U R \cong L.$

(b) $\mathrm{Ann}\,(\mathrm{Ann}\,(L_0; R); L) = L_0,$
$\mathrm{Ann}\,(\mathrm{Ann}\,(R_0; L); R) = R_0,$

for any submodule L_0 of L and any submodule R_0 of R.

COROLLARY 1.2. *Under the assumption of Theorem 1.1 L/L_0 and $\mathrm{Ann}\,(L_0; R)$, as well as $\mathrm{Ann}\,(R_0; L)$ and R/R_0, form an orthogonal pair with respect to U, and each of these modules is isomorphic to the U-character module of the other in the respective pair.*

PROOF. First we assume that L has a composition series, say

$$0 = L_1 \subset L_2 \subset \cdots \subset L_r = L,$$

as a left A-module. Then we have

$$\mathrm{Ann}\,(L_i; R) \supsetneqq \mathrm{Ann}\,(L_{i+1}; R), \quad \text{for } i = 1, 2, \ldots, r-1,$$
$$\mathrm{Ann}\,(L_1; R) = R, \quad \mathrm{Ann}\,(L_r; R) = 0.$$

If, for any element $\bar{x}$ of L_{i+1}/L_i and any element $\bar{y}$ of $\mathrm{Ann}\,(L_i)/\mathrm{Ann}\,(L_{i+1})$, we put

$$(\bar{x}, \bar{y}) = (x, y), \quad (x \in L_{i+1}, \ y \in \mathrm{Ann}\,(L_i)),$$

this pairing does not depend on the choice of representatives x and y. Moreover, if $(\bar{x}, \bar{y}) = 0$ for all $x \in L_{i+1}/L_i$, then we have $\bar{y} = 0$ as an element of $\mathrm{Ann}\,(L_i)/\mathrm{Ann}\,(L_{i+1})$. Therefore $\mathrm{Ann}\,(L_i)/\mathrm{Ann}\,(L_{i+1})$ is isomorphic to a submodule of $\mathrm{Char}_U\,(L_{i+1}/L_i)$, and consequently, since $\mathrm{Char}_U\,(L_{i+1}/L_i)$ is simple, $\mathrm{Ann}\,(L_i)/\mathrm{Ann}\,(L_{i+1})$ is zero or simple. Thus if we denote the length of a composition series of a left (or right) A-module M by $[M]_l$ (or $[M]_r$), we have

$$[R]_r \leq [L]_l. \tag{1}$$

The above argument shows at the same time that

$$[\mathrm{Char}_U L]_r \leq [L]_l, \tag{2}$$

since, if we put $(x, \varphi) = \varphi(x)$ for $x \in L$, $\varphi \in \mathrm{Char}_U L$, L and $\mathrm{Char}_U L$ form a pair and $\mathrm{Ann}\,(L; \mathrm{Char}_U L) = 0$, $\mathrm{Ann}\,(0; \mathrm{Char}_U L) = \mathrm{Char}_U L$.

By symmetry of L and R we have further

$$[L]_l \leq [R]_r. \tag{3}$$

On the other hand, since $\mathrm{Ann}\,(L; R) = 0$, R is isomorphic to a submodule of $\mathrm{Char}_U L$, and hence

$$[R]_r \leq [\mathrm{Char}_U L]_r. \tag{4}$$

Therefore we have by (2) to (4),

$$[L]_l \leq [R]_r \leq [\mathrm{Char}_U L]_r \leq [L]_l.$$

This shows that R is isomorphic to $\mathrm{Char}_U L$, and (a) is proved hereby.

At the same time we see that $\mathrm{Ann}(L_i)/\mathrm{Ann}(L_{i+1})$ is simple, and

$$\mathrm{Ann}\big(\mathrm{Ann}(L_{i+1})\big)/\mathrm{Ann}\big(\mathrm{Ann}(L_i)\big)$$

is also simple. Therefore

$$0 = \mathrm{Ann}\big(\mathrm{Ann}(L_1)\big) \subset \cdots \subset \mathrm{Ann}\big(\mathrm{Ann}(L_r)\big) = L$$

is a composition series of L and $L_i \subseteqq \mathrm{Ann}\big(\mathrm{Ann}(L_i)\big)$. Thus we have

$$L_i = \mathrm{Ann}\big(\mathrm{Ann}(L_i)\big), \quad i = 1, 2, \ldots, r.$$

Since any submodule L_0 of L is a member of some composition series of L we have (b) for any L_0.

PROOF OF COROLLARY 1.2. For $x \in L$, $\bar{x} \in L/L_0$, $y \in \mathrm{Ann}(L_0)$ we put $(\bar{x}, y) = (x, y)$. Then this defines an orthogonal pairing of L/L_0 and $\mathrm{Ann}(L_0; R)$ to U by (b) of Theorem 1.1.

COROLLARY 1.3. *Under the assumption of Theorem* 1.1, *any U-character of a left A-submodule L_0 of L can be extended to a U-character of L.*

PROOF. Since $\mathrm{Char}_U L_0 = R/\mathrm{Ann}(L_0; R)$, any U-character of L_0 is represented as the residue class of some U-character of L. This proves our corollary.

LEMMA 1.4. *Let A be a ring with a unit. Suppose that U satisfies the maximal condition for left and right submodules. Then the following two conditions are equivalent:*

(a) *The U-character module of every simple left A-module as well as that of every simple right A-module is simple.*

(b) *The duality relations*

$$L \cong \mathrm{Char}_U(\mathrm{Char}_U L), \quad R \cong \mathrm{Char}_U(\mathrm{Char}_U R) \tag{5}$$

hold for every simple left A-module L and for every simple right A-module R.

PROOF. Let L be a simple left A-module. For $x \in L$, $\varphi \in \mathrm{Char}_U L$ we shall put

$$(x, \varphi) = \varphi(x),$$

and

$$L_0 = \mathrm{Ann}(\mathrm{Char}_U L; L).$$

Suppose that $L_0 = L$. Then we have $(x, \varphi) = 0$ for every $x \in L$, $\varphi \in \mathrm{Char}_U L$ and hence $\mathrm{Char}_U L = 0$.

(a) $\rightarrow$ (b). In case (a) holds, $\mathrm{Char}_U L$ is simple and L_0 must be zero. Therefore L and $\mathrm{Char}_U L$ form an orthogonal pair and (b) follows at once from Theorem 1.1.

(b) $\rightarrow$ (a). If (b) holds, $\mathrm{Char}_U L \neq 0$, since otherwise we would have $L = 0$ by (b). Hence L and $\mathrm{Char}_U L = R$ form an orthogonal pair.

If R/R_1 is simple for some submodule R_1, then if we put $L_1=\mathrm{Ann}(R_1; L)$, we have either $L_1=L$ or $L_1=0$.

If $L_1=0$, then we have $\mathrm{Char}_U(R/R_1)=0$, since $L_1\cong\mathrm{Char}_U(R/R_1)$. This shows that $R/R_1=0$ by the duality relation in (b). Therefore $L_1=L$. Hence we have $(x,\varphi)=0$ for $x\in L$, $\varphi\in R_1$ and consequently $R_1=0$. Thus R is simple.

REMARK. From the condition (a) only, we cannot conclude in general the validity of (5) for every left A-module possessing a composition series; If A is a ring satisfying the minimum condition and its radical N is not zero, we have $\mathrm{Char}_U A\cong A/N$, $\mathrm{Char}_U(A/N)\cong A/N$, for $U=A/N$, and (a) holds for this A and U.

§2. Quasi-Frobenius rings

In case A is a principal ideal domain $\mathfrak{o}$ (commutative or non-commutative), if we take the residue class module $K/\mathfrak{o}$ as U where K is a quotient field of $\mathfrak{o}$, then the first condition of Theorem 1.1 is satisfied and Theorem 1.1 may be applied; $\mathrm{Char}_{K/\mathfrak{o}}(\mathfrak{o}/\mathfrak{o}a)\cong\mathfrak{o}/a\mathfrak{o}$.

For a commutative ring A we can prove more generally that if A is an integral domain $\mathfrak{o}$ such that (i) $\mathfrak{o}$ is integrally closed in its quotient field K, (ii) every prime ideal of $\mathfrak{o}$ is maximal and (iii) A satisfies the maximum condition for ideals, then the condition of Theorem 1.1 is satisfied for $U=K/\mathfrak{o}$. Indeed, if $\mathfrak{p}$ is a maximal ideal of $\mathfrak{o}$, then $\mathrm{Char}_U(\mathfrak{o}/\mathfrak{p})\cong\mathfrak{p}^{-1}/\mathfrak{o}$ and $\mathfrak{p}^{-1}/\mathfrak{o}$ is simple, since $\mathfrak{p}\,\mathfrak{p}^{-1}=\mathfrak{o}$.

Here our concern lies in the case where A is a ring satisfying the minimum condition. Hereafter we shall always deal with such a ring. In this case we put $U=A$ throughout. First we shall prove

THEOREM 2.1. *Let A be a ring satisfying the minimum condition for left and right ideals and possessing a unit element. Then the following statements are equivalent:*

(a) *The A-character module of every simple left A-module as well as that of every simple right A-module is simple.*

(b) *For every simple left A-module L and for every simple right A-module R, we have*

$$L\cong\mathrm{Char}_A(\mathrm{Char}_A L),\qquad R\cong\mathrm{Char}_A(\mathrm{Char}_A R).$$

(c) *A is quasi-Frobenius.*

PROOF. (a) $\rightleftharpoons$ (b) is proved by Lemma 1.4.

(c) $\rightarrow$ (a). As is shown in NAKAYAMA ([7], p. 9) we have $r(N)=l(N)$ where N is the radical of A, $r(S)=\{x\,|\,ax=0$ for every $a\in S\}$, $l(S)=\{x\,|\,xa=0$ for every $a\in S\}$ for $S<A$. If e is a primitive idempotent of A, then the A-character module of Ae/Ne is isomorphic to $er(N)=el(N)$, and $el(N)$ is simple by the definition of quasi-Frobenius rings.

(a) $\rightarrow$ (c). As above, the A-character module of Ae/Ne is isomorphic to $er(N)$. By (a), $er(N)$ is a simple right A-module and hence $er(N)<l(N)$. Therefore we have $r(N)=l(N)$. These facts, together with (b), imply (c).

Now let A be a quasi-Frobenius ring. If we denote by $A^{(n)}$ and $B^{(n)}$ a free left A-module and a free right A-module with n generators:

$$A^{(n)} = Au_1 \oplus Au_2 \oplus \cdots \oplus Au_n, \qquad B^{(n)} = v_1 A \oplus v_2 A \oplus \cdots \oplus v_n A,$$

then $A^{(n)}$ and $B^{(n)}$ form an orthogonal pair with respect to A; this is seen by putting

$$\Big(\sum_{i=1}^{n} a_i u_i, \sum_{i=1}^{n} v_i b_i\Big) = \sum_{i=1}^{n} a_i b_i \in A.$$

Let L be any finitely generated left A-module. Then it is well known that $L \cong A^{(n)}/L_0$ with a suitable submodule L_0 of $A^{(n)}$ for some n. Therefore by Corollary 1.2 and Theorem 1.1 we obtain immediately the following theorem.

THEOREM 2.2. *Let A be a quasi-Frobenius ring and let L be a finitely generated left A-module. If we put*

$$(x, \varphi) = \varphi(x) \quad \text{for} \quad x \in L,\ \varphi \in \mathrm{Char}_A L,$$

this pairing is orthogonal with respect to A, and $L \cong \mathrm{Char}_A(\mathrm{Char}_A L)$ and the annihilator relations

$$\mathrm{Ann}\big(\mathrm{Ann}(L_0;\ \mathrm{Char}_A L);\ L\big) = L_0$$
$$\mathrm{Ann}\big(\mathrm{Ann}(R_0;\ L);\ \mathrm{Char}_A L\big) = R_0$$

hold for every submodule L_0 of L and every submodule R_0 of $\mathrm{Char}_A L$.

Thus we have an analogy to PONTRJAGIN's theory of character groups of locally compact commutative groups [*10*].

COROLLARY 2.3. *A is quasi-Frobenius if and only if the annihilator relations*

$$l\big(r(L)\big) = L, \qquad r\big(l(R)\big) = R, \tag{6}$$

hold for every left ideal L and every right ideal R of A.

PROOF. The "only if" part is a special case of Theorem 2.2.

To prove the "if" part, let e be a primitive idempotent. Then we have $\mathrm{Char}_A(Ae/Ne) \cong er(N)$, $er(N) = r\big(A(1-e)+Ne\big)$ and $r\big(A(1-e)+Ne\big)$ is simple by (6) since $A(1-e)+Ne$ is a maximal left ideal. Thus the "if" part follows from Theorem 2.1.

REMARK. If we take a finite free A-module $A^{(n)}$ as L, we have a theorem of M. HALL. In [*8*] AZUMAYA has derived Theorem 2.2 from HALL's theorem. Our method yields a simultaneous deduction of Theorem 2.2, Hall's Theorem and NAKAYAMA's characterization of quasi-Frobenius rings (cf. Corollary 2.3 above); indeed, our proof of Theorem 1.1 is an elaboration of NAKAYAMA'S argument in the proof of [*7*], Theorem 6.

Concerning Frobenius rings we have

THEOREM 2.4. *In order that a ring A with a unit be a Frobenius ring it is necessary and sufficient that*

$$\mathrm{Char}_A(A/N) \cong A/N$$

hold for the left A-module A/N and the right A-module A/N, where N is the radical of A.

THEOREM 2.5. *Let A be a quasi-Frobenius ring. Then A is a Frobenius ring if and only if*

$$\mathrm{Char}_A(Z)_l \cong (Z)_r, \qquad \mathrm{Char}_A(Z)_r \cong (Z)_l$$

for every simple two-sided ideal Z, and A is uniserial if and only if the above relations hold for every two-sided ideal Z. Here by $(Z)_l$ and $(Z)_r$ we denote the left ideal Z and the right ideal Z respectively.

We omit the proof.

§3. Submodules of a free module

Let A be a ring satisfying the minimum condition and possessing a unit element.

Let $A^{(n)}$ be a free left A-module: $A^{(n)} = Au_1 \oplus Au_2 \oplus \cdots \oplus Au_n$.

Our purpose in this section is to prove

THEOREM 3.1. *Let L and L' be two A-submodules of $A^{(n)}$ and let ϑ be an isomorphism of $A^{(n)}/L$ onto $A^{(n)}/L'$. Then there exists an isomorphism φ of $A^{(n)}$ onto itself which induces the given isomorphism ϑ and, in particular, L and L' are isomorphic to each other.*

The same holds for right modules.

We shall begin with

LEMMA 3.2. *Let A be a semisimple ring and L a left A-submodule of $A^{(n)}$. Then for a given direct sum decomposition of $A^{(n)}/L$:*

$$A^{(n)}/L = A\bar{v}_1 \oplus A\bar{v}_2 \oplus \cdots \oplus A\bar{v}_n$$

with elements $v_i \in A^{(n)}$, where the bar indicates the residue class, there exist elements w_i of $A^{(n)}$ such that

$$A^{(n)} = Aw_1 \oplus Aw_2 \oplus \cdots \oplus Aw_n,$$

$$w_i \equiv v_i \pmod{L}, \qquad i = 1, 2, \ldots, n.$$

PROOF. Let us denote by $A(1-e_i)$ the set of annihilators of $\bar{v}_i$:

$$A(1-e_i) = \{x \mid x\bar{v}_i = 0,\ x \in A\},$$

where e_i is an idempotent of A; the existence of such an e_i is assured by the semisimplicity of A. Then, if we put

$$v_i' = e_i v_i,$$

we have $v_i' \equiv v_i \pmod{L}$, $v_i' = e_i v_i'$. We now prove that the sum: $Av_1' + Av_2' + \cdots + Av_n' + L$ is direct. If $\sum_i a_i v_i' + x_0 = 0$ for some $a_i \in A$, $x_0 \in L$, then $\sum a_i \bar{v}_i = 0$, and hence $a_i \bar{v}_i' = a_i e_i \bar{v}_i = 0$, and consequently $a_i e_i \in A(1-e_i)$; this shows that $a_i e_i = 0$, and hence $a_i v_i' = 0$. Therefore we have a direct sum decomposition of $A^{(n)}$: $A^{(n)} = Av_1' \oplus Av_2' \oplus \cdots \oplus Av_n' \oplus L$, $v_i' \equiv v_i \pmod{L}$.

Since $Av_i' \cong Ae_i$ and A is semisimple, L is decomposed into a direct sum of L_i $(i=1, 2, \ldots, n)$ such that

$$L_i = Az_i, \quad i = 1, 2, \ldots, n$$
$$\{x \mid xz_i = 0,\ x \in A\} = Ae_i, \quad (1-e_i)z_i = z_i.$$

If we put $w_i = v_i' + z_i$, then we have $Aw_i = Av_i \oplus Az_i \cong A$, since $v_i' = e_i w_i$ and $z_i = (1-e_i)w_i$.

On the other hand, $w_i \equiv v_i'$ (mod L) and hence $w_i \equiv v_i$ (mod L). Since $Aw_i \cong A$, we have

$$A^{(n)} = Aw_1 \oplus Aw_2 \oplus \cdots \oplus Aw_n.$$

Proof of Theorem 3.1. I. Let us first consider the case where A is semisimple. Then there exist n elements v_i of $A^{(n)}$ by Lemma 3.2 such that

$$A^{(n)} = Av_1 \oplus Av_2 \oplus \cdots \oplus Av_n, \quad A^{(n)}/L = A\bar{v}_1 \oplus A\bar{v}_2 \oplus \cdots \oplus A\bar{v}_n.$$

If we put $\vartheta(\bar{v}_i) = \bar{w}_i$, we have

$$A^{(n)}/L' = A\bar{w}_1 \oplus A\bar{w}_2 \oplus \cdots \oplus A\bar{w}_n.$$

By Lemma 3.2 again we can find n elements w_i' of $A^{(n)}$ such that

$$A^{(n)} = Aw_1' \oplus Aw_2' \oplus \cdots \oplus Aw_n', \quad w_i \equiv w_i' \text{ (mod } L).$$

If we define a homomorphism φ by

$$\varphi(\sum a_i v_i) = \sum a_i w_i',$$

then φ is an isomorphism of $A^{(n)}$ onto itself and induces ϑ.

II. Next let N be the radical of A and $N^{(n)}$ the submodule of $A^{(n)}$ consisting of all the elements of the form $\sum a_i w_i$ with $a_i \in N$, $w_i \in A^{(n)}$. We shall first remark that ϑ maps $(N^{(n)}+L)/L$ onto $(N^{(n)}+L')/L'$ isomorphically. Therefore ϑ induces the isomorphism

$$\vartheta^*: \quad A^{(n)}/(N^{(n)}+L) \cong A^{(n)}/(N^{(n)}+L'). \tag{5}$$

On the other hand $A^{(n)}/(N^{(n)}+L) \cong A^{(n)}/N^{(n)}/(N^{(n)}+L)/N^{(n)}$ and $A^{(n)}/N^{(n)}$ is isomorphic to a direct sum of $A\bar{u}_i$, $i=1, 2, \ldots, n$, each of which is isomorphic to the left A-module A/N, where $\bar{u}_i$ means the residue class of u_i modulo $N^{(n)}$. Hence by the first part of the proof we can find an isomorphism φ_0 of $A^{(n)}/N^{(n)}$ onto itself such that φ_0 induces the isomorphism ϑ^* in (5).

If we put $\varphi_0(\bar{u}_i) = \bar{v}_i$, then we have

$$A^{(n)} = Av_1 \oplus Av_2 \oplus \cdots \oplus Av_n, \quad v_i \in \bar{v}_i, \quad i = 1, 2, \ldots, n. \tag{6}$$

To prove this, we first observe that there exist $a_{ij} \in A$, $r_{ij} \in N$ such that $u_i = \sum_j a_{ij} v_j + \sum_j r_{ij} u_j$, since $A^{(n)}/N^{(n)} = A\bar{v}_1 \oplus A\bar{v}_2 \oplus \cdots \oplus A\bar{v}_n$, $N^{(n)} = Nu_1 \oplus Nu_2 \oplus \cdots \oplus Nu_n$.

Therefore we have

$$(E-(r_{ij}))\begin{pmatrix}u_1\\u_2\\\vdots\\u_n\end{pmatrix}=(a_{ij})\begin{pmatrix}v_1\\v_2\\\vdots\\v_n\end{pmatrix},$$

where E means the unit matrix of degree n. Since the radical N is nilpotent, there exists a positive integer k such that $N^k=0$. Then, since $r_{ij}\in N$, we have $(r_{ij})^k=0$ and hence $(E+(r_{ij})+(r_{ij})^2+\cdots+(r_{ij})^{k-1})(E-(r_{ij}))=E$. Therefore we have the expression

$$\begin{pmatrix}u_1\\u_2\\\vdots\\u_n\end{pmatrix}=(a'_{ij})\begin{pmatrix}v_1\\v_2\\\vdots\\v_n\end{pmatrix},\qquad a'_{ij}\in A.$$

Thus we have $A^{(n)}=Av_1+Av_2+\cdots+Av_n$. By the consideration of the length of composition series this is the direct-sum decomposition. Thus if we put

$$\varphi_1(u_i)=v_i,\quad i=1,2,\ldots,n,$$

then φ_1 defines an isomorphism of $A^{(n)}$ onto itself such that φ_1 induces the isomorphism ϑ^* [cf. (5)].

If we denote by $\tilde{w}$ the residue class of $w\in A^{(n)}$ modulo L or L', then the residue class of $\vartheta(\tilde{u}_i)$ modulo $(L'+N^{(n)})/L'$ coincides with the residue class of $\tilde{v}_i$ modulo $(L'+N^{(n)})/L'$. Therefore we can find elements r_i of $N^{(n)}$ such that $\vartheta(\tilde{u}_i)=\tilde{v}_i+\tilde{r}_i$, $i=1,2,\ldots,n$. Hence if we put

$$\varphi(u_i)=w_i,\quad w_i=v_i+r_i,\quad i=1,2,\ldots,n,$$

then we have the direct-sum decomposition:

$$A^{(n)}=Aw_1\oplus Aw_2\oplus\cdots\oplus Aw_n,$$

since $v_i\equiv w_i \bmod N^{(n)}$ [cf. the above proof of (6)]. Hence φ defines an isomorphism of $A^{(n)}$ onto itself which induces the given isomorphism ϑ. This completes our proof.

REMARK. Theorem 3.1 cannot be brought into a sharper form such that $M/L\cong M/L'$ implies $L\cong L'$ for submodules L, L' of an A-module M where M may not be free. It can be shown that this proposition holds for any M if and only if A is semisimple; indeed if $N\neq 0$, there exists some primitive idempotent e with $Ne\neq 0$, and hence if we construct $M=M_1\oplus M_2$, $M_1\cong Ae/Ne$, $M_2\cong Ae$, and put $L=M_1+NM_2$, $L'=M_2$, then $M/L\cong M/L'$, but $L\not\cong L'$.

§ 4. A characterization of quasi-Frobenius rings

We shall first prove the following theorem[4]).

THEOREM 4.1. *Let A be a quasi-Frobenius ring. Then for an A-submodule L of a finite free left A-module $A^{(n)}$ any homomorphism ϑ of L into $A^{(n)}$ can be extended to a homomorphism Θ of $A^{(n)}$ into itself.*

[4]) IKEDA [3] has proved this theorem for $n=1$.

PROOF. Let $A^{(n)} = Au_1 \oplus Au_2 \oplus \cdots \oplus Au_n$ and let $\vartheta(x) = \sum_i \vartheta_i(x)\, u_i$. Then each $\vartheta_i\colon L \to A$ is a left A-character of L. Therefore, in view of Theorem 2.1, it follows from Corollary 1.3 that ϑ_i can be extended to an A-character Θ_i of $A^{(n)}$. Then $\Theta(x) = \sum_i \Theta_i(x) u_i$ clearly defines a homomorphism of $A^{(n)}$ into itself such that $\Theta(x) = \vartheta(x)$ for $x \in L$.

Now we shall establish the following theorem[4]).

THEOREM 4.2. *Let A be a quasi-Frobenius ring. Then for an A-submodule L_1 of a finite free left A-module $A^{(n)}$, any isomorphism ϑ of L_1 onto an A-submodule L_2 of $A^{(n)}$ can always be extended to an isomorphism Θ of $A^{(n)}$ onto itself.*

PROOF. Let $A^{(n)}$ and $B^{(n)}$ be a free left and a free right A-module with generators $\{u_i\}$ and $\{v_i\}$ respectively:

$$A^{(n)} = Au_1 \oplus Au_2 \oplus \cdots \oplus Au_n, \qquad B^{(n)} = v_1 A \oplus v_2 A \oplus \cdots \oplus v_n A.$$

We shall define a pairing of $A^{(n)}$ and $B^{(n)}$ to A by

$$(\sum a_i u_i, \sum v_j b_j) = \sum a_i b_i \in A.$$

Then $A^{(n)}$ and $B^{(n)}$ form an orthogonal pair with respect to A. Since A is quasi-Frobenius, we have

$$\mathrm{Char}_A L_i = B^{(n)}/\mathrm{Ann}(L_i; B^{(n)}), \qquad i = 1, 2,$$

for two submodules L_i of A.

Hence if we define a homomorphism $\lambda\colon B^{(n)}/\mathrm{Ann}(L_2; B^{(n)}) \to B^{(n)}/\mathrm{Ann}(L_1; B^{(n)})$ by

$$(x, \lambda(\bar{y})) = (\vartheta(x), \bar{y}), \qquad x \in L_1,$$

for $\bar{y} \in B^{(n)}/\mathrm{Ann}(L_2, B^{(n)})$, then λ is an isomorphism onto, since ϑ is an isomorphism onto by assumption.

By Theorem 3.1, there exists an isomorphism Λ of $B^{(n)}$ onto itself which maps $\mathrm{Ann}(L_2; B^{(n)})$ onto $\mathrm{Ann}(L_1; B^{(n)})$ and induces the isomorphism λ. We shall define again a homomorphism $\Theta\colon A^{(n)} \to A^{(n)}$ by

$$(\Theta(x), y) = (x, \Lambda(y)) \quad \text{for } x \in A^{(n)},\ y \in B^{(n)}.$$

Then Θ is an isomorphism onto. For $x \in L_1$, $y \in \mathrm{Ann}(L_2, B^{(n)})$ we have $(x, \Lambda(y)) = 0$ and hence $\Theta(x) \in L_2$, since $L_2 = \mathrm{Ann}(\mathrm{Ann}(L_2; B^{(n)}); A^{(n)})$. Therefore we have

$$(\Theta(x), \bar{y}) = (x, \lambda(\bar{y})) \quad \text{for } x \in L_1,$$

and hence $\Theta(x) = \vartheta(x)$ for every $x \in L_1$. This proves the theorem.

COROLLARY 4.3. *Let A be a quasi-Frobenius ring. An isomorphism between two left submodules L_1 and L_2 of a free left A-module $A^{(n)}$ with n generators implies an isomorphism between two left residue class modules $A^{(n)}/L_1$ and $A^{(n)}/L_2$ of $A^{(n)}$. The same holds for right submodules of a finite free right A-module $B^{(n)}$.*

PROOF. It is obvious from Theorem 4.2.

THEOREM 4.4. *Let A be a ring with a unit element which satisfies the minimum condition for left and right ideals and let A possess the following property (P):*

If two left A-submodules L_1 and L_2 of a finite free left A-module $A^{(n)}$ are isomorphic to each other, then two residue class modules $A^{(n)}/L_1$ and $A^{(n)}/L_2$ are isomorphic to each other; similarly the same holds for right submodules of a finite free right A-module $B^{(n)}$ (where $n=1, 2, \ldots$).

Then A is a quasi-Frobenius ring.

PROOF. Let us denote the unit element of A by 1 and the radical of A by N; then we have two direct decompositions of A:

$$A=\sum_{\varkappa=1}^{m}\Big(\sum_{i=1}^{f(\varkappa)} A e_{\varkappa,i}\Big)=\sum_{\varkappa=1}^{m}\Big(\sum_{i=1}^{f(\varkappa)} e_{\varkappa,i} A\Big),$$

where $e_{\varkappa,i}$ are primitive mutually orthogonal idempotents, $1=\sum_{\varkappa=1}^{m}\sum_{i=1}^{f(\varkappa)} e_{\varkappa,i}$ and $Ae_{\varkappa,1}\cong Ae_{\varkappa,i}$, $Ae_{\varkappa,i}\not\cong Ae_{\lambda,j}$ for $\varkappa\neq\lambda$. We put $e_\varkappa=e_{\varkappa,1}$, $E_\varkappa=\sum_i e_{\varkappa,i}$.

If we assume (P) for left simple ideals of A and if two simple left ideals L_1 and L_2 are isomorphic to each other, then by Theorem 3.1 there exists an element a of A such that $L_2=L_1 a$. Hence applying Lemma 2 of IKEDA ([3], p. 204), we obtain that there exists a permutation π of $(1, 2, \ldots, m)$ such that the largest completely reducible left subideal of $Ae_\varkappa$ is a direct sum of simple ideals, each of which is isomorphic to $Ae_{\pi(\varkappa)}/Ne_{\pi(\varkappa)}$ and $r(N)E_\varkappa=E_{\pi(\varkappa)}r(N)\subseteqq l(N)$.

Hence, from the property (P) for right and left simple ideals of A, we have the permutation π of $(1, 2, \ldots, m)$ and $r(N)E_\varkappa=E_{\pi(\varkappa)}l(N)$.

Suppose $e_{\pi(\varkappa)}l(N)\cong t\,\bar{e}_\varkappa\bar{A}$, $t>1$, where $\bar{e}_\varkappa\bar{A}\cong e_\varkappa A/e_\varkappa N$. Then there exist t elements c_j of $e_{\pi(\varkappa)}l(N)\,e_\varkappa$ such that

$$e_{\pi(\varkappa)}l(N)=\sum_{j=1}^{t} e_{\pi(\varkappa)}c_j e_\varkappa A.$$

Let $A^{(2)}=Au_1\oplus Au_2$ be a free left A-module with generators u_1 and u_2. Then $Ae_\varkappa u_1$ and $Ae_\varkappa u_2$ are, by the correspondence $1\leftrightarrow u_1\leftrightarrow u_2$, both isomorphic to the left ideal $Ae_\varkappa$. We construct a left A-module M_0 which is a direct sum of $Ae_\varkappa u_1$ and $Ae_\varkappa u_2$ and introduce into M_0 such a relation that an element $a_1e_\varkappa u_1$ of $Ae_\varkappa u_1$ and an element $a_2e_\varkappa u_2$ of $Ae_\varkappa u_2$ are identified if and only if there exists an element d of A satisfying $a_1e_\varkappa=dc_1e_\varkappa$ and $a_2e_\varkappa=dc_2e_\varkappa$; then we have a module M. It can be proved without much difficulty[5]) that M is a directly indecomposable left A-module.

Put $L_1=Ae_{\pi(\varkappa)}(c_1e_\varkappa u_1-c_2e_\varkappa u_2)$ and $L_2=Ae_{\pi(\varkappa)}c_1e_\varkappa u_1$; then clearly M is isomorphic to a residue class module M_0/L_1 and another residue class module M_0/L_2 is directly decomposable. Thus L_1 and L_2 are isomorphic to $A\bar{e}_{\pi(\varkappa)}$

[5]) See KOETHE's proof of his theorem ([5], p. 40); in our case it should be noted that $r(N)=l(N)$.

and hence to each other, but the residue class modules $(Au_1 \oplus Au_2)/L_1$ and $(Au_1 \oplus Au_2)/L_2$ of $(Au_1 \oplus Au_2)$ $(\cong A^{(2)})$ cannot be isomorphic to each other. This contradicts our property (P) for simple left modules.

Hence $e_{\pi(\varkappa)}l(N)$ is simple and similarly $r(N)e_\varkappa$ is also simple. This completes our proof.

REMARK. In order that $L \cong L'$ implies $M/L \cong M/L'$ for submodules L, L' of any left A-module M where M may not be free, it is necessary and sufficent that A be semisimple; indeed if $N \neq 0$, there exists some primitive idempotent e with $Ne \neq 0$ and hence if we construct $M = M_1 \oplus M_2$, $M_1 \cong Ae$, $M_2 \cong Ne$, and put $L = NM_1$, $L' = M_2$, then $L \cong L'$ but $M/L \ncong M/L'$.

From Corollary 4.3 and Theorem 4.4 we obtain

THEOREM 4.5. *Let A be a ring satisfying the minimum condition and possessing a unit element. A is quasi-Frobenius if and only if A possesses property (P).*

§5. (M_u)-Modules of quasi-Frobenius rings

Let L be a finite (= finitely generated) left A-module. L is called an (M_u)-module according to GASCHÜTZ [1], if the following condition is satisfied: For any finite left A-module M containing L as its A-submodule there exists some A-submodule L_0 such that M is the direct sum of L and L_0: $M = L \oplus L_0$. In case A is an algebra over a field, NAGAO and NAKAYAMA ([6], p. 169) have determined the structure of (M_u)-modules.

We shall first establish

THEOREM 5.1. *Let A be a quasi-Frobenius ring. A finite left A-module L is an (M_u)-module if and only if L is a direct sum of A-submodules each of which is isomorphic to some indecomposable direct summand $Ae_{\varkappa,i}$ of A* (for the meaning of $e_{\varkappa,i}$, cf. § 4).

To prove this theorem it is convenient to prove the following theorem.

THEOREM 5.2. *Let A be a quasi-Frobenius ring. Then any finite left A-module is isomorphic to a submodule of a finite free left A-module $A^{(n)}$ for suitable n.*

PROOF. For a left A-module L we take the character module of L. Then $\mathrm{Char}_A L$ is isomorphic to a residue class A-right module $B^{(n)}/R$. By Theorem 2.1 and Corollary 1.2 we have a left submodule L_1 of $A^{(n)}$ which is isomorphic to the $\mathrm{Char}_A B^{(n)}/R$. Thus L is isomorphic to L_1, because L is isomorphic to $\mathrm{Char}_A (B^{(n)}/R) \cong \mathrm{Char}_A (\mathrm{Char}_A L)$.

PROOF OF THEOREM 5.1. Let L be an (M_u)-module. Then by Theorem 5.2 we may assume that L is contained in a finite free A-module $A^{(n)}$ with some n. Since, by definition, L is a direct summand of $A^{(n)}$, we see at once that L has the structure asserted in the theorem.

Conversely, let L be a finite left A-module which is a direct sum of A-submodules each isomorphic to some $Ae_{\varkappa,i}$. Let M be any finite A-module containing L as its A-submodule. By Theorem 5.2 we can assume that M is contained in a finite free A-module $A^{(m)}$ with some m. Then by the

assumption concerning the structure of L there exists a direct-sum decomposition of $A^{(n)}$ with a suitable n ($\geq m$), such that

$$A^{(n)} = P \oplus Q, \quad \text{with } P \cong L.$$

Since by Theorem 4.2 the isomorphism $P \cong L$ can be extended to an isomorphism of $A^{(n)}$ onto itself, we have another direct-sum decomposition of $A^{(n)}$: $A^{(n)} = L \oplus Q_0$. Since $L \subseteq M \subseteq A^{(n)}$, we obtain at once the direct-sum decomposition of M: $M = L \oplus (M \cap Q_0)$. Thus L is an (M_u)-module.

Combining our theorem with a theorem of NAGAO-NAKAYAMA [6] we have

COROLLARY 5.3. *In case A is a quasi-Frobenius ring, the notions of (M_u)-modules and (M_0)-modules are identical.*

THEOREM 5.4. *Let A be a ring which contains a unit element and satisfies the minimum condition. Then A is quasi-Frobenius if and only if A is an (M_u)-module when considered as a left and a right A-module.*

PROOF. The "only if" part is obvious from Theorem 5.1. We shall prove the "if" part. First, assume that some right ideal $e_\lambda r(N)$ is neither simple nor zero. Then $e_\lambda r(N)$ has a composition series, say

$$e_\lambda r(N) = e_\lambda M_1 \supset e_\lambda M_2 \supset \cdots \supset e_\lambda M_t \supset (0),$$

where $t \geq 2$. If $e_\lambda M_1 / e_\lambda M_2 \cong \bar{e}_\varkappa \bar{A}$, $e_\lambda M_2 / e_\lambda M_3 \cong \bar{e}_\varrho \bar{A}$, where $\varkappa$ is not necessarily different from ϱ, then there exist elements c_1 and c_2 of $er(N)$ such that $c_1 = e_\lambda c_1 e_\varkappa \in e_\lambda M_1$, $c_2 = e_\lambda c_2 e_\varrho \in e_\lambda M_2$ and $c_1 \notin e_\lambda M_2$. Let $A^{(2)} = Au_1 \oplus Au_2$ be a free left A-module with generators u_1 and u_2. Since $Ae_\lambda c_1 \cong Ae_\lambda c_2 \cong \bar{A}\bar{e}_\lambda$, $Ae_\lambda c_1 e_\varkappa u_1$ and $Ae_\lambda c_2 e_\varrho u_2$ are, by the correspondence $e_\lambda c_1 e_\varkappa u_1 \leftrightarrow e_\lambda c_2 e_\varrho u_2 \leftrightarrow \bar{e}_\lambda$, both isomorphic to the left simple ideal $\bar{A}\bar{e}_\lambda$. Now, we identify the submodules $Ac_1 e_\varkappa u_1$ and $Ac_2 e_\varrho u_2$ of Au_1 and $Ae_\varrho u_2$, and thus construct a new left module T:

$$T = (Au_1, Ae_\varrho u_2), \qquad Au_1 \cap Ae_\varrho u_2 = Ac_1 u_1 = Ac_2 u_2.$$

Then T clearly contains $Au_1 \cong A$. But Au_1 cannot appear in any direct decomposition of T as a direct summand as will be proved below. Thus we arrive at a contradiction to our assumption that A is a (M_u)-module as a left A-module. Hence we obtain that if A is a (M_u)-module as a left A-module then every $e_\lambda r(N)$ is either a simple right ideal or zero. Therefore $e_\lambda r(N) \subseteq e_\lambda l(N)$, and consequently $r(N) \subseteq l(N)$. Similarly if A is an (M_u)-module as a right A-module, then $l(N)e_\varkappa \subseteq r(N)e_\varkappa$, $\varkappa = 1, 2, \ldots, m$, and hence $l(N) \subseteq r(N)$. Therefore we have $r(N) = l(N)$, and $e_\lambda r(N) = e_\lambda l(N)$, $\lambda = 1, 2, \ldots, m$, are simple right ideals and $r(N)e_\lambda = l(N)e_\lambda$, are simple left ideals. This shows that A is quasi-Frobenius.

To complete our proof it remains to show that Au_1 cannot appear in any direct decomposition of T as a direct summand. Assume that $T = Au_1 \oplus Q$, for some submodule Q of T. Then we have

$$e_\varrho u_2 = e_\varrho y u_1 + e_\varrho q, \qquad \text{where } y \in A,\ q \in Q.$$

Hence $Q \supseteq A(-e_\varrho y u_1 + e_\varrho u_2)$. Here we consider an A-left homomorphism σ of Ae_ϱ onto $A(-e_\varrho y u_1 + e_\varrho u_2)$ defined by

$$\sigma: \quad a e_\varrho \to a(-e_\varrho y u_1 + e_\varrho u_2).$$

If the kernel H of σ is not zero, there exists a non-zero element h of H such that $-he_\varrho y u_1 + he_\varrho u_2 = 0$, and from the definition there exists an element a of A satisfying

$$he_\varrho y = ac_1 e_\varkappa, \quad \text{and} \quad he_\varrho = ac_2.$$

Since $Aac_2 = Ac_2$, we have $bhe_\varrho = bac_2 e_\varrho = c_2 e_\varrho$ and $bhe_\varrho y e_\varkappa = bac_1 e_\varkappa$ for a suitable element b of A. Hence $e_\lambda bae_\lambda c_2 = c_2 \neq 0$. Therefore $e_\lambda bae_\lambda$ is not contained in $e_\lambda N e_\lambda$, since $c_2 \in r(N)$. Hence there exists an inverse element $e_\lambda f e_\lambda$ of $e_\lambda bae_\lambda$ in $e_\lambda A e_\lambda$. Therefore we obtain

$$e_\lambda f e_\lambda bhe_\varrho y e_\varkappa = c_1, \quad \text{and} \quad e_\lambda f e_\lambda bhe_\varrho = c_2.$$

Hence $c_2 e_\varrho y e_\varkappa = c_1$, and this shows $c_1 \in e_\lambda M_\varkappa$, but it is a contradiction. This contradiction is caused from assumption $H \neq 0$. Hence σ must be an isomorphism. Hence $[Q]_l \geq [Ae_\varrho]_l$. From the assumption that T is a direct sum of Au_1 and Q, we have

$$[T]_l = [Au_1]_l + [Q]_l \geq [A]_l + [Ae_\varrho]_l.$$

On the other hand, by the definition of T,

$$[T]_l < [A]_l + [Ae_\varrho]_l.$$

This is a contradiction.

COROLLARY 5.5. *Let A be a ring possessing a unit element and satisfying the minimum condition. Then A is quasi-Frobenius if and only if all directly indecomposable summands $Ae_\varkappa$ and $e_\lambda A$, $\varkappa, \lambda = 1, 2, \ldots, m$ of A are (M_u)-modules.*

§ 6. A representation-theoretical characterization of quasi-Frobenius algebras

Let A be an algebra of a finite rank over a ground field F with a unit element, and let $S^{(n)}$ be a representation in F which is given by a representation module isomorphic to a finite free left A-module $A^{(n)}$. If we have two splittings of $S^{(n)}$ into constituents

$$S^{(n)} = \begin{pmatrix} H_1 & 0 \\ * & K_1 \end{pmatrix}, \qquad S^{(n)} = \begin{pmatrix} H_2 & 0 \\ * & K_2 \end{pmatrix},$$

then by Theorem 3.1 the two constituents K_1 and K_2 are equivalent as representations of A in F if H_1 is equivalent to H_2, and by Theorem 4.5 the converse holds when A is a quasi-Frobenius algebra.

For any representation H we can take a suitable representation $S^{(n)}$ so that H may appear at the left-upper corner of a reduced form of $S^{(n)}$:

$$S^{(n)} = \begin{pmatrix} H & 0 \\ * & K \end{pmatrix},$$

but it is not always possible to find $S^{(n)}$ with

$$S^{(n)} = \begin{pmatrix} K' & 0 \\ * & H \end{pmatrix}.$$

Now we shall prove

THEOREM 6.1. *Let A be an algebra of a finite rank over a ground field F with a unit element. If every representation H of A appears at the right-lower corner of a splitting form of $S^{(n)}$ with a suitable n:*

$$S^{(n)} = \begin{pmatrix} K & 0 \\ * & H \end{pmatrix},$$

then A is quasi-Frobenius, and conversely.

PROOF. We shall prove the first part of the theorem. Let H be a representation which is given by a primitive right ideal $e_\varkappa A$ as a representation module; then the dual module $(e_\varkappa A)^*$ of $e_\varkappa A$ is, by assumption, contained in a finite free left A-module $A^{(n)}$. By a theorem of NAGAO and NAKAYAMA ([6], p. 169), the left A-module $(e_\varkappa A)^*$ is an (M_u)-module. Hence, there exists a primitive left ideal Ae_λ such that $Ae_\lambda \cong (e_\varkappa A)^*$. This shows that A is a quasi-Frobenius algebra.

The converse is obtained from Theorem 5.2.

References

[1] GASCHÜTZ, W.: Über den Fundamentalsatz von MASCHKE zur Darstellungstheorie der endlichen Gruppen. Math. Z. **56**, 376—387 (1952). — [2] GASCHÜTZ, W.: Über modulare Darstellungen endlicher Gruppen, die von freien Gruppen induziert werden. Math. Z. **60**, 274—286 (1954). — [3] IKEDA, M.: A characterization of quasi-Frobenius ring. Osaka Math. J. **4**, 203—209 (1952). — [4] YUTAKA KAWADA: A note on isomorphism and similarity of ideals (forthcoming in J. Math. Soc. Japan). — [5] KOETHE, G.: Verallgemeinerte Abelsche Gruppen mit hyperkomplexen Operatorring. Math. Z. **39**, 31—44 (1934). — [6] NAGAO, H., and T. NAKAYAMA: On the structure of (M_0)- and (M_u)-modules. Math. Z. **59**, 164—170 (1953). — [7] NAKAYAMA, T.: On Frobeniusean algebra II. Ann. of Math. **42**, 1—21 (1941). — [8] NAKAYAMA, T., and G. AZUMAYA: Algebra II (in Japanese). Tokyo 1954. — [9] NESBITT, C. J., and R. M. THRALL: Some ring theorems with applications to modular representation. Ann. of Math. **47**, 551—567 (1946). — [10] PONTRJAGIN, L.: Topologialc groups. Princeton Univ. Press 1939. — [11] THRALL, R. M.: Some generalizations of quasi-Frobenius algebras. Trans. Amer. Math Soc. **64**, 173—183 (1948).

Dept. of Math., Tokyo University of Education, Tokyo (Japan)

(Eingegangen am 14. Februar 1956)

SCHÖNEBORN, H.
Math. Zeitschr. Bd. 65, S. 429—441 (1956)

Über den Zusammenhang zwischen Dualitäts- und Vollständigkeitseigenschaften bei gewissen topologischen Abelschen Gruppen

Von

HEINZ SCHÖNEBORN

In der vorliegenden Note werden die vom Verf. in zwei vorangegangenen Arbeiten [3] und ([4][1])) betrachteten topologischen, primären, Abelschen Gruppen unter zusätzlichen Abzählbarkeitsvoraussetzungen im Hinblick auf ihre (Cauchy-)Vollständigkeits- und ihre Dualitätseigenschaften (bei Charakterengruppenbildung) untersucht. Es wird gezeigt, wie sich der Begriff der Vollständigkeit und gewisse in früheren Untersuchungen in Verbindung damit auftretende Begriffsbildungen in durchsichtiger Weise dualisieren und zueinander in Beziehung setzen lassen und ferner, daß zwischen diesen Vollständigkeitseigenschaften und der Eigenschaft, selbstdual[2]) zu sein, ein einfacher Zusammenhang besteht. Insbesondere wird dadurch auch die Frage nach der Gültigkeit des PONTRJAGINschen Dualitätssatzes für die betrachteten Gruppen völlig und in befriedigender Weise gelöst, indem einfache und wohlbekannte Vollständigkeitseigenschaften als notwendige und hinreichende Kriterien dafür gewonnen werden[3]).

In Nr. 1 werden die Begriffe der „Abzählbarartigkeit" und „schwachen Abzählbarartigkeit" als Verallgemeinerungen des ersten HAUSDORFFschen Abzählbarkeitsaxioms eingeführt. Ihre Bedeutung für unsere Betrachtungen beruht darauf, daß sie (im Gegensatz zum ersten HAUSDORFFschen Abzählbarkeitsaxiom) bei Übergang zur Charakterengruppe erhalten bleiben. Unter die „schwach abzählbarartigen" Gruppen fallen insbesondere z. B. die lokalkompakten Gruppen (ohne irgendeine Abzählbarkeitsvoraussetzung). — In Nr. 2 wird zur Eigenschaft der (Cauchy-)Vollständigkeit die sich später als

[1]) In Zukunft zitiert mit Sch. I bzw. Sch. II.

[2]) Eine Gruppe A heißt „selbstdual", wenn der PONTRJAGINsche Dualitätssatz — im folgenden auch kurz als „Dualitätssatz" bezeichnet — für A gilt.

[3]) Eine ohne Abzählbarkeitsvoraussetzungen gültige notwendige und hinreichende Bedingung für die Richtigkeit des Dualitätssatzes für eine torsionstopologische zyklovollständige, primäre Abelsche Gruppe (vgl. den Schluß der Einleitung) ergibt sich sofort aus Sch. II Satz 7 in Verbindung mit einem von H. LEPTIN hergeleiteten Satz ([1] Satz S. 261). Man hat hierbei nur zu beachten, daß unter Voraussetzung der für die Gültigkeit des Dualitätssatzes notwendigen Bedingung (K_1) aus Sch. II Satz 6 die von LEPTIN in [1] zugrunde gelegte Topologisierung der Charakterengruppen mit der üblichen und vom Verf. verwendeten übereinstimmt. — Für den hier behandelten Sonderfall abzählbarartiger Gruppen ergeben sich — wie gezeigt wird [vgl. unten (3.1) und anschließende Bemerkungen] — andere und unter Umständen handlichere Kriterien.

dazu dual erweisende der „Antivollständigkeit" definiert und beide Vollständigkeitseigenschaften werden in Beziehung gesetzt zu gewissen in früheren Untersuchungen (Sch. II und [1]) über den Dualitätssatz aufgetretenen Bedingungen. — Nr. 3 enthält als Hauptresultat den Satz (3.1), welcher unter anderem die Äquivalenz von Vollständigkeit und Antivollständigkeit einerseits und Dualitätssatz andererseits aussagt. — In Nr. 4 und Nr. 5 wird die (vorher schon hervorgetretene) Dualität einiger wichtiger Begriffe wie „vollständig" und „antivollständig", „präkompakt" und „prädiskret" usw. genauer gefaßt und näher untersucht und auf einige naheliegende offengebliebene Fragen und Verallgemeinerungsmöglichkeiten hingewiesen.

Alle im folgenden betrachteten Gruppen sind, auch wenn nicht ausdrücklich hervorgehoben, additiv geschriebene „torsionstopologische, zyklovollständige, primäre, Abelsche Gruppen" im Sinne von Sch. I Nr. 1—4, d.h. topologische Abelsche Gruppen A, deren Topologie den folgenden Bedingungen genügt:

1. Es gibt ein aus offenen Untergruppen A_α von A bestehendes vollständiges System von Nullumgebungen in A, derart, daß jedes von Null verschiedene $a \in A$ von Primzahlpotenzordnung modulo jedem A_α ist.

2. Die in A abgeschlossenen Hüllen der (gewöhnlichen) zyklischen Untergruppen von A sind (Cauchy-)vollständig.

Wie in Sch. I gezeigt wurde, sind dann die in A abgeschlossenen Hüllen der zyklischen Untergruppen von A topologisch isomorph entweder zur Additionsgruppe $\mathfrak{R}_p$ des Ringes der ganz-p-adischen Zahlen (p = ganz-rationale Primzahl) oder zu gewöhnlichen zyklischen p-Gruppen und A besitzt den Ring der ganz-p-adischen Zahlen als „natürlichen Operatorenbereich".

Die Menge aller offenen Untergruppen von A bildet einen „*Filter*" im Sinne von Sch. I Nr. 1, den sog. „*charakteristischen Filter*" von A. Hinsichtlich der Grundbegriffe „Filterbasis", „topologischer Homomorphismus", „TYCHONOFF-Summe" usw. s. gleichfalls Sch. I Nr. 1.

1. A sei eine zyklovollständige, Abelsche, primäre topologische Gruppe.

Definition. A heißt „*abzählbarartig*", wenn mindestens eine der beiden folgenden Bedingungen erfüllt ist:

1. Der charakteristische Filter von A besitzt eine aus abzählbar vielen offenen Untergruppen A_i bestehende Basis $A = A_0 \supset A_1 \supset A_2 \supset \cdots$ (1. HAUSDORFFsches Abzählbarkeitsaxiom).

2. Es gibt eine abgezählte aufsteigende Kette $K_0 = 0 \subset K_1 \subset K_2 \subset \cdots$ kompakter Untergruppen K_i in A, so daß jede kompakte Untermenge von A in (mindestens) einer der Untergruppen K_i enthalten ist. — Genauer heißt A „*abzählbarartig von 1. Art*" bzw. „*abzählbarartig von 2. Art*", je nachdem die erste oder die zweite der obigen Bedingungen erfüllt ist.

Versteht man — wie üblich — unter einer Gruppe vom Typus p^∞ eine diskrete Torsionsgruppe mit abzählbar vielen Erzeugenden a_i mit den Relationen $p \cdot a_1 = 0$, $p \cdot a_2 = a_1$, $p \cdot a_3 = a_2$, ... (p ganz-rationale Primzahl), dann bildet die topologische direkte Summe („TYCHONOFF-Summe"; vgl. Sch. I S. 459) abzählbar vieler Gruppen vom Typus p^∞ ein Beispiel für eine abzähl-

barartige Gruppe von 1. Art, welche nicht von 2. Art ist. Sind umgekehrt $\mathfrak{R}^+_{p,i}$ $(i=1,2,\ldots)$ Gruppen vom Typus der (kompakten) Additionsgruppe der ganz-p-adischen Zahlen, A die gewöhnliche direkte Summe der $\mathfrak{R}^+_{p,i}$ mit der Topologie, die durch den charakteristischen Filter aller (operatorzulässigen) Untergruppen A_α von A, deren Restklassengruppe A/A_α eine Torsionsgruppe ist, definiert wird, so ist A abzählbarartig von 2. Art, nicht aber von 1. Art. — Ebenso ist jede diskrete Gruppe abzählbarartig von 1. Art, während jede kompakte Gruppe abzählbarartig von 2. Art ist.

Die Frage, unter welchen Voraussetzungen eine abzählbarartige Gruppe gleichzeitig abzählbarartig von 1. und von 2. Art ist, wird beantwortet durch

(1.1) *Eine Gruppe A ist genau dann abzählbarartig gleichzeitig von 1. und von 2. Art, wenn sie lokalkompakt ist und dem 2.* HAUSDORFF*schen Abzählbarkeitsaxiom genügt.*

Beweis. Sei A lokalkompakt und genüge dem 2. Abzählbarkeitsaxiom. Dann ist A natürlich abzählbarartig von 1. Art und besitzt eine offene, kompakte, dem 2. Abzählbarkeitsaxiom genügende Untergruppe K_1. A/K_1 ist diskret und abzählbar und durch sukzessive Hinzunahme je endlich vieler mod K_1 verschiedener Elemente von A zu K_1 konstruiert man der Reihe nach eine aufsteigende Kette $K_1 < K_2 < \cdots$ kompakter Gruppen, hinsichtlich deren A abzählbarartig von 2. Art ist. — Umgekehrt: Sei A abzählbarartig von 1. und von 2. Art, $A_1 > A_2 > \cdots$ ein absteigendes vollständiges System offener Untergruppen, $K_1 < K_2 < \cdots$ ein aufsteigendes System kompakter Untergruppen, hinsichtlich dessen A abzählbarartig von 2. Art ist. Natürlich genügen die K_i dem 2. Abzählbarkeitsaxiom. Wäre keine der Gruppen K_i offen, d.h. enthielte keines der K_i irgendeine der Gruppen A_j, so gäbe es zu jedem Paar i, j $(i, j = 1, 2, \ldots)$ ein $a_{ij} \in A$ mit $a_{ij} \in A_j$, $a_{ij} \notin K_i$. Die Folge $a_{11}, a_{22}, \ldots$ ist eine Nullfolge, also, wenn man Null hinzunimmt, eine kompakte Menge und diese läge nach Konstruktion in keinem der K_i — im Widerspruch zu der Abzählbarartigkeit 2. Art von A. Mithin ist eines der K_i, etwa K_1, offen und A lokalkompakt. Die diskrete Restklassengruppe A/K_1 ist, da K_1 kompakt, mit A offenbar abzählbarartig 2. Art, also abzählbar. Und hieraus folgt, daß A dem 2. Abzählbarkeitsaxiom genügt.

Definition. Die Gruppe A heißt *„schwach abzählbarartig von 1. Art"* bzw. *„schwach abzählbarartig von 2. Art"*, wenn die erste bzw. die zweite der beiden folgenden Bedingungen erfüllt ist:

1. A enthält eine kompakte Untergruppe K, so daß A/K abzählbarartig von 1. Art ist.

2. A enthält eine offene Untergruppe A', welche abzählbarartig von 2. Art ist.

Offenbar ist jede abzählbarartige Gruppe auch schwach abzählbarartig von der gleichen Art, während eine beliebige lokalkompakte Gruppe nicht notwendig abzählbarartig von 1. oder 2. Art, wohl aber stets schwach abzählbarartig von 1. und von 2. Art ist.

(1.2) *Die Gruppe A ist genau dann schwach abzählbarartig von 1. und von 2. Art, wenn sie lokalkompakt ist.*

Es bleibt nur zu beweisen: Ist A schwach abzählbarartig von 1. und von 2. Art, dann ist A lokalkompakt. In der Tat: Sei K kompakt, so daß A/K abzählbarartig von 1. Art und A' offen und abzählbarartig von 2. Art. Dann ist $A'/A' \cap K = (A'+K)/K$ abzählbarartig von 1. Art und mit A' und wegen der Kompaktheit von $A' \cap K$ gleichzeitig abzählbarartig von 2. Art. Nach (1.1) ist also $A'/A' \cap K$ lokalkompakt und genügt dem 2. Abzählbarkeitsaxiom, woraus die lokale Kompaktheit von A' und damit die von A folgt.

Da sich die Sätze der folgenden Paragraphen von abzählbarartigen Gruppen fast mühelos auf schwach abzählbarartige Gruppen übertragen lassen, beschränken wir uns der Einfachheit und Kürze halber im folgenden auf die auf abzählbarartige Gruppe bezüglichen Betrachtungen.

2. (2.1) *Eine abzählbarartige Gruppe A von 1. Art ist dann und nur dann vollständig, wenn die folgende Bedingung* (K) *erfüllt ist:*

(K) *Jede abgeschlossene Untergruppe von A mit kompaktem Erzeugendensystem ist kompakt*[4]).

Eine abzählbarartige Gruppe 2. Art genügt der Bedingung (K) *immer, ist aber nicht notwendig vollständig.*

Beweis. Ist A vollständig, so gilt (K) nach Sch. II Satz 1a. — Genüge nun A der Bedingung (K) und sei $a_1, a_2, \ldots$ eine Cauchy-Folge aus A. Dann ist $\lim (a_i - a_{i+1}) = 0$ und die Elemente $a_1, (a_1 - a_2), (a_2 - a_3), \ldots$ bilden zusammen mit der 0 eine kompakte Untermenge von A, erzeugen wegen (K) also eine kompakte Untergruppe K. Diese enthält alle Elemente a_i und mithin existiert in K und erst recht in $A \supseteq K$ der $\lim a_i$, woraus die Vollständigkeit von A folgt. — Für abzählbarartige Gruppen 2. Art folgt (K) aus der Definition. Ein einfaches Beispiel für eine nicht vollständige abzählbarartige Gruppe 2. Art bildet die von H. LEPTIN [2] § 6 angegebene Gruppe: Seien $A_1, A_2, \ldots$ abzählbar viele zyklische Gruppen von der gleichen Primzahlordnung p, $A = \sum A_i$ die gewöhnliche direkte Summe der A_i. Topologisiert man A, indem man sämtliche Untergruppen von endlichem Index in A für offen erklärt, dann ist A präkompakt[5]), aber nicht kompakt, also nicht vollständig und außerdem erweisen sich sämtliche kompakten Untermengen von A als endlich. Mithin ist A eine nicht vollständige abzählbarartige Gruppe 2. Art.

Dem Begriff der Vollständigkeit einer Gruppe A stellen wir den — wie sich zeigen wird — dazu weitgehend dualen der Antivollständigkeit gegenüber.

Definition. Die (nicht notwendig abzählbarartige) Gruppe A heißt „*antivollständig*", wenn gilt: Ist A' eine beliebige Untergruppe von A derart, daß

[4]) Identisch mit der in Sch. II S. 24 formulierten und dort mit (K_1) bezeichneten Bedingung.

[5]) Eine Gruppe heißt präkompakt, wenn ihre vollständige Hülle kompakt ist. Die torsionstopologische Gruppe A ist genau dann präkompakt, wenn sämtliche offene Untergruppen von endlichem Index in A sind.

für jede kompakte Untergruppe $K \subseteq A$ der Durchschnitt $A' \cap K$ offen in K ist, dann ist A' offen in A.

Als Gegenstück zu (2.1) gilt jetzt:

(2.2) *Eine abzählbarartige Gruppe A von 2. Art ist dann und nur dann antivollständig, wenn die folgende Bedingung* (L) *erfüllt ist:*

(L) *Ist A' eine abgeschlossene Untergruppe von A derart, daß alle kompakten Untermengen von A/A' endlich sind, dann ist A' offen in A.*

Eine abzählbarartige Gruppe 1. Art genügt der Bedienung (L) immer, ist aber nicht notwendig antivollständig.

Beweis. A sei abzählbarartig von 2. Art und antivollständig, A' eine abgeschlossene Untergruppe von A der in (L) genannten Art. Dann ist für jede kompakte Untergruppe K von A sicher $K/A' \cap K$ endlich, also $A' \cap K$ offen in K und wegen der Antivollständigkeit ist A' offen in A, A/A' diskret — woraus (L) folgt. Genüge nun die von 2. Art abzählbarartige Gruppe A der Bedingung (L) und sei A' eine beliebige Untergruppe von A derart, daß für alle kompakten Untergruppen $K_1 < K_2 < \cdots$ von A, hinsichtlich deren A abzählbarartig von 2. Art ist, der Durchschnitt $K_i \cap A'$ offen in K_i ist. Dann existiert zu jedem K_i eine offene Untergruppe A_i von A mit $K_i \cap A_i = K_i \cap A'$. Da die K_i eine aufsteigende Folge bilden, gilt jetzt für jedes $m \geq 1$:

$$K_m \cap A_1 \cap A_2 \cap \cdots \cap A_m \subseteq K_m \cap A'$$

und (natürlich) offen in K_m, $K_m \cap A_{m+1} \cap A_{m+2} \cap \cdots = K_m \cap A'$, also $K_m \cap \left(\bigcap_i A_i\right) \subseteq K_m \cap A'$ und offen in K_m. Für $B = \bigcap A_i$ ist also $B \subseteq A'$ und in A/B ist jede kompakte Untermenge endlich. Nach (L) ist also B und mithin auch A' offen in A — woraus die Antivollständigkeit von A folgt. — Sei nun A abzählbarartig von 1. Art und A' eine abgeschlossene Untergruppe der in (L) genannten Art. Wäre A' nicht offen in A, so gäbe es eine Nullfolge $a_1, a_2, \ldots$ $(a_i \in A)$ mit $a_i \notin A'$ und $\bar{A} = A/A'$ enthielte im Widerspruch zur Definition von A' die unendliche kompakte Untermenge $\bar{0}, \bar{a}_1, \bar{a}_2, \ldots$. Mithin ist (L) erfüllt. — Daß aber eine von 1. Art abzählbarartige Gruppe A nicht antivollständig zu sein braucht, zeigt das Beispiel der gewöhnlichen direkten Summe $A = \sum_{i=1}^{\infty} A_i$, wo die A_i Gruppen der festen Primzahlordnung p sind, wenn in A die Untergruppe von der Form $A^{(j)} = \sum_{i \neq j} A_i$ für offen erklärt werden. In A ist jede kompakte Untergruppe endlich, aber die 0 ist nicht offen in A.

Die Bedeutung der Bedingung (L) wird noch verdeutlicht und in Beziehung zu der vom Verf. in Sch. II S. 29 aufgestellten Forderung (K_2) gebracht durch

(2.3) *A sei abzählbarartig. Dann gilt* (L) *genau dann, wenn die folgende Bedingung* (K_2) *erfüllt ist:*

(K_2) *Ist M eine Untermenge von A, $a \in A$ ein Häufungspunkt von M, dann gibt es eine kompakte Untermenge N von A derart, daß a bereits Häufungspunkt von $N \cap M$ ist.*

Dem Beweis von (2.3) schicken wir den folgenden Hilfssatz voran:

(2.4) *Genügt A der Bedingung* (K), *dann ist* (K_2) *äquivalent zu jeder der beiden folgenden Bedingungen:*

(a) *Ist M eine Untermenge von A, $a \in A$ ein Häufungspunkt von M, dann gibt es eine kompakte Untergruppe K von A, so daß a Häufungspunkt von $M \cap K$ ist.*

(b) *Ist für eine Untermenge M und für jede kompakte Untergruppe K von A stets $M \cap K$ offen in K, dann ist M offen in A.*

Beweis von (2.4). Die Äquivalenz von (a) und (K_2) unter der Voraussetzung (K) ist klar. Wir beweisen die Äquivalenz von (a) und (b). Aus (a) folgt (b): Sei M' das Komplement von M in A und $M' \cap K$ abgeschlossen in K für jede kompakte Untergruppe K von A. $a \in A$ sei ein Häufungspunkt von M'. Wegen (a) existiert eine kompakte Untergruppe K' von A, so daß a Häufungspunkt von $M' \cap K'$, also $a \in M' \cap K' \subseteq M'$ ist, d.h. M' ist abgeschlossen, also M offen in A. — Aus (b) folgt (a): Sei $a \in A$ Häufungspunkt von $M \subseteq A$, $\overline{M}$ die in A abgeschlossene Hülle von M, M' die Menge aller Elemente von $\overline{M}$ mit Ausnahme von a. Wäre für keine kompakte Untergruppe K von A das Element a Häufungspunkt von $K \cap M$, so wäre $M' \cap K$ abgeschlossen in K für jedes K, d.h. für das (a enthaltende) Komplement M'' von M' in A wäre $M'' \cap K$ in K offen für jedes K, also wegen (b) M'' offen in A und $a \in M''$ kann nicht Häufungspunkt von M sein.

Beweis von (2.3). Ist A abzählbarartig von 1. Art, dann ist (L) nach (2.2) immer erfüllt und ebenso gilt stets (K_2), da, wenn $a \in A$ Häufungspunkt von $M \subseteq A$ ist, eine gegen a konvergente Folge aus M existiert. Im Falle der abzählbarartigen Gruppen 1. Art ist also nichts zu beweisen. — Sei nun A abzählbarartig von 2. Art. Dann ist (L) eine Folge von (K_2). Sei nämlich A' eine abgeschlossene Untergruppe von A derart, daß in $\bar{A} = A/A'$ jede kompakte Untermenge endlich ist. f sei die kanonische Abbildung von A auf $\bar{A}$, $M \subseteq A$ und $a \in A$ ein Häufungspunkt von M. Dann ist $\bar{a} = f(a)$ entweder in $\overline{M} = f(M)$ enthalten oder Häufungspunkt von $\overline{M}$. Wegen (2.4) (a) existiert eine kompakte Untergruppe K von A, so daß a Häufungspunkt von $K \cap M$ ist. Aus der Endlichkeit von $f(K \cap M)$ folgt $\bar{a} \in f(K \cap M) \subseteq \overline{M}$ d.h. aus „a Häufungspunkt von M" folgt, daß die a enthaltende Nebenschar mod A' ein Element von M enthalten muß. Die zu diesem Resultat führenden Voraussetzungen und Schlüsse bleiben — wie leicht zu sehen — erhalten, wenn A' durch den Durchschnitt $A' \cap A_\alpha$, wo A_α eine beliebige offene Untergruppe von A ist, ersetzt wird. Es ergibt sich, daß sich die durch die Zuordnung der Häufungspunkte definierte Topologie von A nicht ändert, wenn A' zum System der offenen Untergruppen von A hinzugenommen wird, d.h. A' ist offen.

Sei nun umgekehrt (L) erfüllt und also nach (2.2) A antivollständig. $K_1 < K_2 < \cdots$ sei eine Folge kompakter Gruppen, hinsichtlich deren A abzählbarartig von 2. Art ist. Ist nun $M \subseteq A$ mit $0 \in M$ und $M \cap K_i$ offen in K_i für alle i, so existiert zunächst in $M \cap K_1$ eine in K_1 offene Untergruppe B_1'.

9272-65 28 B

B_1' ist eine in $M \cap K_2$ enthaltene abgeschlossene Untergruppe von K_2 und Durchschnitt aller B_1' enthaltenden offenen Untergruppen $K_2^{(j)}$ von K_2. Aus $\bigcap_j K_2^{(j)} = B_1' \subseteqq K_2 \cap M$ und den Offenheits- bzw. Abgeschlossenheitseigenschaften dieser Mengen sowie der Kompaktheit von K_2 folgt jetzt leicht die Existenz einer in K_2 offenen Untergruppe B_2' mit $B_1' \subseteqq B_2' \subseteqq K_2 \cap M$ usw. Ist B die Vereinigungsmenge der B_i', so ist B eine Untergruppe von A mit $B \subseteqq M$ und $B \cap K_i \supseteqq B_i'$ offen in K_i und wegen der Antivollständigkeit ist B offen in A. M enthält also mit 0 eine in A offene Umgebung der 0 und genau so ergibt sich das entsprechende für jedes Element von M. Daraus folgt, daß M offen in A ist und nach (2.4) folgt (K_2).

3. Die Sätze der vorangegangenen Nummern sowie die insbesondere in Sch. II hergeleiteten Sätze und Begriffsbildungen — vor allem über die Charakterengruppen einer primären, zyklovollständigen, torsionstopologischen, Abelschen Gruppe — erlauben jetzt die Formulierung des folgenden Hauptsatzes:

(3.1) *Sei A abzählbarartig. Dann sind die folgenden Aussagen* (1) *bis* (4) *zueinander äquivalent:*

(1) *A genügt den Bedingungen* (K) *und* (L) *aus* (2.1) *bzw.* (2.2).

(2) *A genügt den Bedingungen* (K) *und* (K_2) *aus* (2.1) *bzw.* (2.3).

(3) *A genügt dem Dualitätssatz, d.h. die (wie üblich gebildete und mit p-adischen Operatoren versehene — vgl. Sch. II Nr. 2 und 3) Charakterengruppe $C[C[A]]$ der Charakterengruppe $C[A]$ von A ist zu A (hinsichtlich der üblichen Abbildung φ von A in $C[C[A]]$ — vgl. Sch. II Nr. 5) topologisch isomorph.*

(4) *A ist vollständig und antivollständig.*

Beweis. Aus (1) folgt (2) und aus (2) folgt (1). Dies ist der Inhalt von (2.3). Aus (2) folgt (3). Dies ergibt sich aus dem Satz 7 von Sch. II. — Aus dem gleichen Satz folgt auch, daß für eine (3) genügende Gruppe die Bedingung (K) erfüllt ist. Wir zeigen, daß auch (L) und damit (1) aus (3) folgt: Ist A abzählbarartig von 1. Art, dann ist (L) nach (2.2) immer erfüllt und nichts weiter zu beweisen. Sei A abzählbarartig von 2. Art und (3) erfüllt. Dann ist $C[A]$ — wie leicht zu sehen — abzählbarartig von 1. Art und genügt ebenfalls dem Dualitätssatz, also auch der Bedingung (K) und ist nach (2.1) vollständig. A' sei eine abgeschlossene Untergruppe von A, so daß in A/A' alle kompakten Untermengen endlich sind. Dann sind in dem Annullator $C_{A'} \subseteqq C[A]$ von A' in $C[A]$ alle offenen Untergruppen von endlichem Index und wegen der Vollständigkeit von $C[A]$ und der Abgeschlossenheit von $C_{A'}$ in $C[A]$ ist auch $C_{A'}$ vollständig, also kompakt. Mithin ist der Annullator $A^{*\prime} \subseteqq C[C[A]]$ von $C_{A'}$ in $C[C[A]]$ offen. $A^{*\prime}$ ist aber genau die A' entsprechende Untergruppe bei der den topologischen Isomorphismus von A auf A^* vermittelnden Abbildung φ und mithin ist A' offen in A. Hieraus folgt (L). — Damit ist gezeigt, daß (1), (2) und (3) äquivalent sind. Sei jetzt (4) erfüllt. Aus (2.1) und (2.2) folgt dann die Gültigkeit von (K) und (L), d.h. (1). — Sei umgekehrt (1) und damit (2) und (3) für A erfüllt. Wir haben

die Vollständigkeit und die Antivollständigkeit von A zu zeigen. Ist zunächst A von 1. Art, dann folgt die Vollständigkeit von A wegen (K) aus (2.1). Ferner ergibt sich die Antivollständigkeit von A wegen der Gültigkeit von (K_2) aus (2.4) (b). Ist aber A von 2. Art, so kann A wegen (3) als Charakterengruppe einer gleichfalls dem Dualitätssatz genügenden Gruppe A^0 von 1. Art angesehen werden. Da A^0 der Bedingung (K_2) genügt, ist A vollständig nach Sch. II S. 29. Die Antivollständigkeit von A ist wegen (L) eine Folge von (2.2). Damit ist alles bewiesen.

Durch (3.1) ist also insbesondere gezeigt, daß die in Sch. II Satz 7 in Verbindung mit (K) als hinreichend für die Gültigkeit des Dualitätssatzes nachgewiesene Bedingung (K_2) im abzählbarartigen Falle auch notwendig ist. Ferner ist durch (3.1) die Frage nach der Gültigkeit des Dualitätssatzes auf die nach der geläufigen und vielfach leicht zu verifizierenden Eigenschaft der Vollständigkeit und die dazu duale der Antivollständigkeit zurückgeführt. Die Frage, inwieweit die Eigenschaften der Vollständigkeit und der Antivollständigkeit [und entsprechend auch die Eigenschaften (K) und (L)] insbesondere im Hinblick auf Charakterengruppenbildung als zueinander „dual" betrachtet werden können, wird in der nächsten Nummer behandelt. — Im folgenden sollen noch einige Eigenschaften der Charakterengruppe $C[A]$ einer abzählbarartigen Gruppe A betrachtet werden:

(3.2) *Ist A abzählbarartig, so ist $C[A]$ (und damit auch jede höhere Charakterengruppe von A) vollständig und antivollständig.*

Beweis. Sei zunächst A von 1. Art. $\hat{A}$ sei die vollständige Hülle von A, $c \in C[A]$ ein Charakter von A. Wir definieren nun einen Charakter $\hat{c}$ von $\hat{A}$ folgendermaßen: Für $a \in A$ setzen wir $\hat{c}(a) = c(a)$ [wo $c(a)$ natürlich das Bild von a in der zur Bildung der Charaktere benutzten Bezugsgruppe Z bedeutet — vgl. Sch. II S. 18f.]. Ist allgemeiner $\hat{a} = \lim a_i \in \hat{A}$ $(a_i \in A, i = 1, 2, \ldots)$, so gilt für fast alle i: $c(a_i) = c(a_{i+1})$ und wir setzen $\hat{c}(\hat{a}) = c(a_i)$ für hinreichend großes i. Dann hängt $\hat{c}(\hat{a})$ unabhängig von der speziellen Wahl der gegen $\hat{a}$ konvergenten Folge der $a_i \in A$ nur von c (und natürlich $\hat{a}$) ab und wir haben so zu jeder Abbildung c von A eindeutig eine Abbildung $\hat{c}$ von $\hat{A}$ definiert und zu verschiedenen c gehören verschiedene $\hat{c}$. Wie leicht zu sehen, ist jedes solche $\hat{c}$ ein Homomorphismus von $\hat{A}$ in die Bezugsgruppe Z und der Kern von $\hat{c}$ in $\hat{A}$ ist einfach die in $\hat{A}$ abgeschlossene Hülle des Kerns des zu $\hat{c}$ gehörigen Homomorphismus c von A in Z. Mithin ist $\hat{c}$ ein Charakter von $\hat{A}$ und jeder Charakter von $\hat{A}$ kann so erhalten werden. Die Charakterengruppen $C[A]$ und $C[\hat{A}]$ sind also hinsichtlich der oben definierten Zuordnung der $c \in C[A]$ zu den $\hat{c} \in C[\hat{A}]$ zunächst algebraisch (Operator-)isomorph. Um zu zeigen, daß $C[A]$ und $C[\hat{A}]$ aber sogar topologisch isomorph sind, genügt es zu beachten: a) Ist $B \subseteq A$ eine Untergruppe mit kompakten Erzeugendensystem von A, so besitzt die in $\hat{A}$ abgeschlossene Hülle $\tilde{B}$ von B gleichfalls ein kompaktes Erzeugendensystem. b) Ist $\hat{B}$ eine Untergruppe mit kompaktem Erzeugendensystem von $\hat{A}$, dann gibt es eine Untergruppe B

von A mit kompaktem Erzeugendensystem, so daß $\hat{B}$ in der in $\hat{A}$ abgeschlossenen Hülle $\tilde{B}$ von B enthalten ist. a) ist fast trivial, während b) im vorliegenden Falle (A genügt dem ersten Abzählbarkeitsaxiom!) in geläufiger Weise durch ein Diagonalverfahren bewiesen werden kann. Aus a) folgt, daß die Topologie von $C[A]$ höchstens feiner ist als die von $C[\hat{A}]$, während zufolge b) die Topologie von $C[A]$ höchstens gröber als die von $C[\hat{A}]$ sein muß. Mithin können $C[A]$ und $C[\hat{A}]$ identifiziert werden und $C[A]$ genügt mit $C[\hat{A}]$ nach (3.1) dem Dualitätssatz, ist also — wieder nach (3.1) — vollständig und antivollständig. — Ist nun A von 2. Art, dann $C[A]$ ersichtlich von 1. Art und man sieht leicht, daß $C[A]$ der Bedingung (K) genügt, also nach (2.1) vollständig und wegen der Gültigkeit von (K_2) nach (2.3) und (3.1) auch antivollständig ist.

4. In dieser Nummer sollen die wichtigsten der bisher aufgetretenen Eigenschaften wie Vollständigkeit, Antivollständigkeit, (K) und (L) im Hinblick auf ihre „Dualität" untersucht werden. Zunächst stellen wir dem in Sch. I definierten Begriff des Filters den des „Antifilters" gegenüber:

Definition. Ein System $\mathfrak{u}$ von Nebenscharen U_μ nach abgeschlossenen Untergruppen B_μ der Gruppe A heißt ein „*Antifilter*" (in A), wenn I. mit den endlich vielen Nebenscharen $U_{\mu_1}, \ldots, U_{\mu_n}$ der Untergruppen $B_{\mu_1}, \ldots, B_{\mu_n}$ auch die die U_{μ_i} enthaltende Nebenschar der abgeschlossenen Vereinigungsgruppe der B_{μ_i} und II. mit U_μ auch jede abgeschlossene in U_μ enthaltene Nebenschar zu $\mathfrak{u}$ gehört.

Ein Teilsystem $\mathfrak{b}_\mathfrak{u}$ von $\mathfrak{u}$ heißt eine *Basis* des Antifilters $\mathfrak{u}$, wenn $\mathfrak{b}_\mathfrak{u}$ der obigen Bedingung I. genügt und jede Schar aus $\mathfrak{u}$ in einer Schar aus $\mathfrak{b}_\mathfrak{u}$ enthalten ist.

Die weitere Dualisierung der auf Filter bezüglichen Grundbegriffe zu solchen, die sich auf Antifilter beziehen, ist — soweit hier von Bedeutung — so leicht, daß auf eine Darstellung im einzelnen verzichtet werden kann.

Unter dem „*charakteristischen Antifilter*" $\mathfrak{s}$ einer Gruppe A verstehen wir den von den kompakten Untergruppen von A gebildeten Antifilter.

Weitere wichtige Antifilter in der Gruppe A sind:

Der Antifilter $\mathfrak{s}'$ aller abgeschlossenen Untergruppen mit kompaktem Erzeugendensystem und der Antifilter $\mathfrak{s}''$ aller abgeschlossenen präkompakten Untergruppen von A.

Analog stellen wir dem aus allen offenen Untergruppen von A bestehenden charakteristischen Filter $\mathfrak{S}$ von A gegenüber:

Den Filter $\mathfrak{S}'$ aller abgeschlossenen Untergruppen A' von A mit der Eigenschaft, daß in A/A' alle kompakten Unter*mengen* endlich sind und den Filter aller abgeschlossenen Untergruppen A'' von A mit der Eigenschaft, daß in A/A'' alle kompakten *Untergruppen* endlich sind.

Offenbar ist $\mathfrak{s} \subseteq \mathfrak{s}' \subseteq \mathfrak{s}''$ und $\mathfrak{S} \subseteq \mathfrak{S}' \subseteq \mathfrak{S}''$.

Definition. Eine Gruppe A heiße „*prädiskret*", wenn der charakteristische Antifilter $\mathfrak{s}$ von A nur aus endlichen Gruppen besteht.

Nun gilt:

(4.1) *Die Gruppe A ist kompakt genau dann, wenn sie 1. präkompakt und 2. vollständig ist.*

Die Gruppe A ist diskret genau dann, wenn sie 1. prädiskret und 2. antivollständig ist.

Der Beweis ist bekannt bzw. trivial.

Offenbar ist die Bedingung (K) äquivalent zu $\mathfrak{s} = \mathfrak{s}'$, während (L) gleichbedeutend mit $\mathfrak{S} = \mathfrak{S}'$ ist. Ferner ist für abzählbarartige Gruppen 1. Art stets $\mathfrak{s}' = \mathfrak{s}''$, während für abzählbarartige Gruppen 2. Art stets $\mathfrak{S}' = \mathfrak{S}''$ gilt. Mit Rücksicht auf (2.1) bzw. (2.2) ergibt sich also:

(4.2) *Eine abzählbarartige Gruppe von 1. bzw. 2. Art ist genau dann vollständig bzw. antivollständig (und damit nach* (3.1) *vollständig und antivollständig), wenn jede abgeschlossene präkompakte Untergruppe kompakt* ($\mathfrak{s} = \mathfrak{s}''$) *bzw. jede prädiskrete Restklassengruppe nach einer abgeschlossenen Untergruppe diskret* ($\mathfrak{S} = \mathfrak{S}''$) *ist.*

Bemerkungen. 1. Eine Gruppe A ist bekanntlich (vgl. Sch. I S. 458 und Satz 15) linear kompakt, wenn A vollständig und für jede offene Untergruppe A_α von A die (diskrete) Restklassengruppe A/A_α endlichrangig (vgl. Sch. I Nr. 8) ist. Dies gibt Anlaß zu folgenden zu den vorangegangenen analogen Begriffsbildungen:

A heiße *linear-präkompakt*, wenn für jede offene Untergruppe A_α von A die Restklassengruppe A/A_α endlichrangig ist.

A heiße *linear-prädiskret*, wenn der charakteristische Antifilter von A nur aus endlichrangigen Gruppen besteht.

Dann ist offenbar eine Gruppe linear kompakt genau dann, wenn sie linear-präkompakt und vollständig ist. Dementsprechend definieren wir weiter:

A heiße *linear-diskret*, wenn sie 1. linear-prädiskret, und 2. antivollständig ist.

Man verifiziert leicht, daß die linear-diskreten Gruppen genau die S-topologischen Gruppen aus Sch. I Nr. 5 sind.

Die Definition von „linear-diskret" weicht formal und inhaltlich von der bei H. LEPTIN [*1*, S. 258] gegebenen ab, scheint aber im Hinblick auf die hier entwickelte Systematik der Begriffe „diskret" und „kompakt" die näherliegende und zweckmäßigere zu sein.

2. Nicht beantwortet werden kann hier die Frage, welche Bedeutung die Bedingung $\mathfrak{S}' = \mathfrak{S}''$ für abzählbarartige Gruppen 1. Art bzw. die Bedingung $\mathfrak{s}' = \mathfrak{s}''$ für abzählbarartige Gruppen 2. Art hat. Eine vollständige Gruppe 1. Art genügt den Bedingungen (K) und (L), ist nach (3.1) also auch antivollständig und daraus folgt leicht $\mathfrak{S} = \mathfrak{S}''$, also erst recht $\mathfrak{S}' = \mathfrak{S}''$. Ebenso folgt für abzählbarartige Gruppen 2. Art $\mathfrak{s} = \mathfrak{s}''$ und damit $\mathfrak{s}' = \mathfrak{s}''$ aus der Antivollständigkeit. Ist aber auch umgekehrt für Gruppen 1. Art die Vollständigkeit eine Folge von $\mathfrak{S}' = \mathfrak{S}''$ und für Gruppen 2. Art die Antivollständigkeit eine Folge von $\mathfrak{s}' = \mathfrak{s}''$? Eine bejahende Antwort auf diese Fragen würde bedeuten, daß für abzählbarartige Gruppen die Bedingungen der

ständigkeit und Antivollständigkeit äquivalent sind. Denn ist A etwa eine abzählbarartige Gruppe 2. Art, so würde aus der Vollständigkeit von A $\mathfrak{s}=\mathfrak{s}''$ und damit $\mathfrak{s}'=\mathfrak{s}''$ und — falls Antivollständigkeit eine Folge von $\mathfrak{s}'=\mathfrak{s}''$ ist — die Antivollständigkeit von A folgen, während umgekehrt eine antivollständige Gruppe 2. Art, da sie (K) und (L) genügt, nach (3.1) auch stets vollständig ist. Analog für Gruppen 1. Art. — In auch für nicht abzählbarartige Gruppen gültiger Form lassen sich die obigen Fragen wie folgt formulieren:

(F_1) Ist stets

a) die Bedingung (K) äquivalent zu $\mathfrak{S}'=\mathfrak{S}''$?

b) die Bedingung (L) äquivalent zu $\mathfrak{s}'=\mathfrak{s}''$?

oder auch allgemeiner:

(F_2) a) Besitzt jede nichtkompakte Gruppe unendliche prädiskrete Restklassengruppen?

b) Enthält jede nichtdiskrete Gruppe unendliche präkompakte Untergruppen?

5. Die in den vorangegangenen Nummern herausgearbeitete Dualität der Eigenschaften (K) und (L) (bzw. von Vollständigkeit und Antivollständigkeit) für abzählbarartige Gruppen findet ihren Ausdruck gleichfalls in der Theorie der Charakterengruppen dieser Gruppen:

Sei A eine abzählbarartige Gruppe, $C[A]$ ihre — wie üblich gebildete — Charakterengruppe. Dann lautet der angekündigte Satz:

(5.1) *a) Die Bedingung* (K) *ist für A genau dann erfüllt, wenn gilt:*

(K$^+$) *Ist C_B eine offene Untergruppe von $C[A]$, dann gibt es eine kompakte Untergruppe B von A, deren Annullator C_B ist. Oder anders gesagt: Die Annullatoren der kompakten Untergruppen von A bilden den ganzen charakteristischen Filter von $C[A]$.*

b) Die Bedingung (L) *ist für A genau dann erfüllt, wenn gilt:*

(L$^+$) *Ist C^B eine kompakte Untergruppe von $C[A]$, dann gibt es eine offene Untergruppe B von A, welche von C^B annulliert wird. Oder anders gesagt: Diejenigen abgeschlossenen Untergruppen von $C[A]$, welche offene Untergruppen von A annullieren, bilden den ganzen charakteristischen Antifilter von $C[A]$.*

Dem Beweis von (5.1) schicken wir voran den folgenden Hilfssatz:

(5.2) *Sei $\mathring{C}[A]$ die nicht topologisierte Charakterengruppe von A, aufgefaßt als Operatorgruppe mit dem Ring der ganz-p-adischen Zahlen als Operatorenbereich* (vgl. Sch. II Nr. 2), *$\mathring{C}^B$ eine (operatorzulässige) Untergruppe von $\mathring{C}[A]$, B die volle von $\mathring{C}^B$ annullierte Untergruppe von A, $\mathring{C}_B \subseteqq \mathring{C}[A]$ der Annullator von B. Ist dann $\mathring{C}_D$ der Annullator einer Untergruppe D von A und ist $\mathring{C}_D \cap \mathring{C}^B$ von endlichem Index in $\mathring{C}^B$, dann ist $\mathring{C}_D \cap \mathring{C}_B$ von endlichem Index in $\mathring{C}_B$.*

Beweis. Indem wir uns von vornherein Restklassengruppenbildung von A nach B vorgenommen denken, können wir ohne Beschränkung der Allgemeinheit $B=0$ $\mathring{C}_D \subseteqq \mathring{C}_B = C[A]$ voraussetzen und hinfort $\mathring{C}^B$ mit $\mathring{C}^0$ be-

zeichnen. $\mathring{C}^{D'} = \mathring{C}_D \cap \mathring{C}^0$ ist von endlichem Index in $\mathring{C}^0$; $D' \subseteq A$ sei die volle von $\mathring{C}^{D'}$ annullierte Untergruppe von A, $\mathring{C}_{D'}$ der Annullator von D'. Offenbar ist $D' \supseteq D$, $\mathring{C}_{D'} \subseteq \mathring{C}_D$, $\mathring{C}_{D'} \supseteq \mathring{C}^{D'}$, $\mathring{C}_{D'} \cap \mathring{C}^0 = \mathring{C}^{D'}$. Die Restklassengruppe $\mathring{C}' = \mathring{C}[A]/\mathring{C}_{D'}$ ist zu identifizieren mit der Charakterengruppe von D'; die den Gruppen $\mathring{C}^0$, $\mathring{C}_D$ bei der Restklassengruppenbildung entsprechenden Gruppen seien $\mathring{C}^{0\prime}$, $\mathring{C}'_D$. $\mathring{C}^{0\prime}$ ist endlich und die volle von $\mathring{C}^{0\prime}$ annullierte Untergruppe von D' ist die Nullgruppe, d.h. D' und damit D sind endlich. Daraus ergibt sich, daß der Index von $\mathring{C}_D$ in $C[A]$ endlich ist — woraus die Behauptung von (5.2) folgt.

Beweis von (5.1). a) Wurde bereits in Sch. II Satz 3 bewiesen. b) Erstens: Sei (L) für A erfüllt, C^B eine kompakte Untergruppe von $C[A]$, B die größte von C^B annullierte Untergruppe von A, $C_B \subseteq C[A]$ der Annullator von B. Nach dem Hilfssatz (5.2) und wegen (3.2) ist C_B kompakt und es ist leicht zu sehen, daß C_B mit der Charakterengruppe $C[A/B]$ von A/B identifiziert werden kann. Wegen der Kompaktheit von $C[A/B]$ ist in A/B jede kompakte Untermenge endlich und mithin wegen (L) B offen in A, woraus (L^+) folgt.

Zweitens: Sei (L^+) erfüllt, B eine abgeschlossene Untergruppe von A derart, daß in A/B jede kompakte Untermenge endlich ist. C_B sei der Annullator von B. Dann kann wieder C_B mit der Charakterengruppe $C[A/B]$ identifiziert werden, und es ist klar, daß C_B kompakt sein muß. Da B die volle von C_B annullierte Untergruppe von A ist, folgt wegen (L^+) die Offenheit von B in A.

Schlußbemerkungen. Einige der in den vorangegangenen Nummern für abzählbarartige Gruppen bewiesenen Sätze bleiben auch für nicht abzählbarartige Gruppen richtig. So z.B. die Äquivalenz der gemeinsamen Gültigkeit der Bedingungen (K) und (L) einerseits und der Gültigkeit des Dualitätssatzes andererseits [(3.1) (1) und (3)], ebenso die Gleichwertigkeit von (K) und (K^+) sowie die von (L) und (L^+) [(5.1)]. Im allgemeinen *nicht* richtig sind hingegen für nicht abzählbarartige Gruppen z.B. die Sätze (2.3) — (L) ist zwar stets eine Folge von (K_2), nicht aber umgekehrt — (3.1) und (3.2). Insbesondere folgt aus Vollständigkeit bzw. Antivollständigkeit zwar auch im nicht abzählbarartigen Falle stets (K) bzw. (L), aber auch aus (K) und (L) zusammen braucht nicht notwendig Vollständigkeit bzw. Antivollständigkeit zu folgen [vgl. (3.1) (1) und (4)]. Trotzdem behalten die Begriffe der Vollständigkeit und Antivollständigkeit auch für nicht abzählbarartige Gruppen ihren Sinn und ihre (duale) Bedeutung, wie schon der folgende Satz zeigt:

Ist A eine (nicht notwendig abzählbarartige) dem Dualitätssatz genügende Gruppe, so ist A (über (L) *hinaus) sogar antivollständig bzw. (über* (K) *hinaus) sogar vollständig genau dann, wenn die Charakterengruppe $C[A]$ vollständig bzw. antivollständig ist.*

Der Beweis ergibt sich aus den folgenden Bemerkungen: Ist A antivollständig, so folgt: Ist f eine Abbildung von A in die Gruppe vom Typus p^∞,

so daß für jede kompakte Untergruppe K von A die von f auf K induzierte Abbildung f_K ein topologischer (d.h. stetiger und offener) Homomorphismus von K ist, dann ist f ein topologischer Homomorphismus von A und hieraus ergibt sich nach einem Satz von KRULL (Sch. II Anm. 3) sofort die Vollständigkeit von $C[A]$. — Sei nun umgekehrt A nicht antivollständig. Dann existiert eine nicht offene Untergruppe $B' \subseteq A$ derart, daß für jede kompakte Untergruppe $K \subseteq A$ der Durchschnitt $B' \cap K$ offen in K ist. Unter Zuhilfenahme des Dualitätssatzes sieht man leicht, daß B' nicht abgeschlossen in A sein kann (betrachte andernfalls die Restklassengruppe A/B' und beachte, daß die Charakterengruppe von A/B' kompakt wird). Dann gibt es sicher eine in A nicht abgeschlossene Obergruppe B von B', für welche die (gewöhnliche Operator-)Restklassengruppe A/B entweder endlich zyklisch oder vom Typus p^∞ ist, und für welche natürlich $B \cap K$ für jede kompakte Untergruppe K von A offen in K ist. Sei f eine (sicherlich existierende) homomorphe Abbildung von A in die Gruppe vom Typus p^∞ mit dem Kern B. Der von f induzierte Homomorphismus f_K von K ist für jede kompakte Untergruppe K von A ein topologischer Homomorphismus. Sei U_K die Menge aller Charaktere von A, welche für K mit der Abbildung f_K identisch sind. U_K ist dann eine Nebenschar der offenen Untergruppe von $C[A]$, welche Annullator von K ist. Durchläuft K alle kompakten Untergruppen von A, so bilden die entsprechenden U_K einen Cauchyfilter F in $C[A]$. Besäße F einen nichtleeren Durchschnitt, so wäre dieser mit der Abbildung f von A identisch. Aber da der Kern B von f nicht abgeschlossen und f mithin kein Charakter von A ist, muß der Durchschnitt der U_K leer sein und $C[A]$ ist nicht vollständig. Mithin ist die Vollständigkeit von $C[A]$ [unter den für A gemachten Voraussetzungen (K) und (L)] gleichbedeutend mit der Antivollständigkeit von A. — Die duale Behauptung, nach welcher die Vollständigkeit von A äquivalent zur Antivollständigkeit von $C[A]$ ist, läßt sich entsprechend beweisen, folgt aber auch bereits aus dem Dualitätssatz auf Grund der Bemerkung, daß A Charakterengruppe von $C[A]$ ist.

Literatur

[1] LEPTIN, H.: Abelsche Gruppen mit kompakten Charaktergruppen und Dualitätstheorie gewisser linear-topologischer Abelscher Gruppen. Abh. Math. Sem. Univ. Hamburg **19**, 244—263 (1954). — [2] LEPTIN, H.: Zur Dualitätstheorie projektiver Limites Abelscher Gruppen. Abh. Math. Sem. Univ. Hamburg **19**, 264—268 (1954). — [3] SCHÖNEBORN, H.: Über gewisse Topologien in Abelschen Gruppen I. Math. Z. **59**, 455—473 (1954). — [4] SCHÖNEBORN, H.: Über gewisse Topologien in Abelschen Gruppen II. Math. Z. **60**, 17—30 (1954).

Bonn, Mathematisches Institut der Universität

(Eingegangen am 26. Februar 1956)

RAMANUJAN, M. S.
Math. Zeitschr. Bd. 64, S. 442—447 (1956)

Theorems on the product of quasi-Hausdorff and Abel transforms

By

M. S. RAMANUJAN

1. INTRODUCTION. O. Szász [5] raised the following question about the "product" of two summability methods for sequences: If a sequence $\{s_n\}$ is summable by a regular method A then is the B-transform of $\{s_n\}$, where B is a regular sequence-to-sequence method, also summable by the method A to the same sum as before? The question was finally answered in the affirmative by Szász [6] himself and PATI [2] in the case when i) A is the Abel method and B is a regular sequence-to-sequence Hausdorff transform (in particular the Cesàro or the Euler transform [5]) and by Szász [6] when ii) A is the Borel Exponential method and B is a regular Hausdorff method and iii) when A is Abel method and B the Hardy-Littlewood-Fekete circle method which is essentially a sequence-to-sequence method ([1], § 9.11). RAJAGOPAL ([3], Theorem I) has extended the results i) and ii) to the case when A is a suitably restricted power series method, known to include the Abel and Borel methods and B is a regular Hausdorff method. I propose now to extend iii) to a transform B, which is a regular quasi-Hausdorff transform known to include the circle method ([1], §§ 11.20, 11.21 (1)) as a particular case. My extension which appears as Theorem 2 of this paper, has an analogue when A, B are series-to-series transformations. This analogue, Theorem 1, is given first as being perhaps the more interesting of the two theorems.

2. NOTATIONS AND LEMMAS. A transformation which converts a series $\sum_n a_n$ to a series $\sum_n b_n$ is said to be *quasi-Hausdorff* when $\{b_n\}$ is defined by

$$(H_1^*, \mu) \quad b_n = \sum_\varkappa \lambda^*_{n\varkappa} a_\varkappa$$

where

$$(1) \qquad \lambda^*_{n\varkappa} = 0 \quad (\varkappa < n), \qquad \lambda^*_{n\varkappa} = \binom{\varkappa}{n} \Delta^{\varkappa - n} \mu_n \quad (\varkappa \geq n).$$

The quasi-Hausdorff series-to-series transformation is a special case of the transformation of $\sum a_n$ to $\sum b_n$ where,

$$(2) \qquad b_n = \sum_{\varkappa=0}^{\infty} h_{n\varkappa} a_\varkappa.$$

If the transformation given by (2) is regular i.e. limit preserving, then the matrix of transformation H is said to be an α-*matrix*. Following VERMES [7], we define an α-matrix each of whose rows has the limit 0, i.e. $\lim_{\varkappa \to \infty} h_{n\varkappa} = 0$ for each n, as an α_0-*matrix*.

On the other hand, a quasi-Hausdorff transformation which converts a sequence s_n $(n=0, 1, 2, \ldots)$ to a sequence t_n $(n=0, 1, \ldots)$ is defined by ([1], § 11.19)

$$(H_2^*, \mu) \quad t_n = \sum_{\varkappa=0}^{\infty} \lambda_{n\varkappa}^* s_\varkappa$$

where $\lambda_{n\varkappa}^*$ is given in terms of μ_n as before.

If in either case, μ_n is a *moment constant of rank n*, generated by a function $\chi(u)$ which is real and of bounded variation for $0 \leq u \leq 1$, we have

$$\mu_n = \int_0^1 u^n \cdot d\chi(u) \tag{3}$$

and we may suppose without loss of generality that $\chi(0)=0$. If in addition,

$$\chi(1) = 1 \quad \text{and} \quad \chi(+0) = \chi(0) = 0 \tag{4}$$

then μ_n is called a *regular moment constant*.

In the proofs of the theorems, I require the following lemmas.

LEMMA 1. *The series-to-series quasi-Hausdorff transformation* (H_1^*, μ) *is regular if, and only if,*

i) μ_n *is a moment constant, i.e. given by* (3) *and* ii) $\mu_0 = 1$ *i.e.* $\int_0^1 d\chi = 1$.

This result has been proved by me elsewhere ([4], Theorem II). Now, since in a quasi-Hausdorff matrix, each row tends to 0 if, and only if, each column of the corresponding Hausdorff matrix tends to 0 which happens if and only if, $\chi(+0)=0$, we have, after Lemma 1,

LEMMA 2. *The quasi-Hausdorff matrix* (H_1^*, μ) *is an* α_0*-matrix if, and only if* μ_n *is a regular moment constant.*

LEMMA 3. *Given an* α_0*-matrix H defining a transformation of the series* $\sum_n a_n$ *to the series* $\sum_n b_n$, *there exists a T-matrix B defined by*

$$b_{n\varkappa} = g_{n\varkappa} - g_{n\cdot\varkappa+1} \quad \text{where} \quad g_{n\varkappa} = h_{0\varkappa} + h_{1\varkappa} + \cdots + h_{n\varkappa} \tag{5}$$

which transforms the sequence $\{s_n\}$ *of partial sums of* $\sum a_n$ *into the sequence* $\{t_n\}$ *of partial sums of* $\sum b_n$ *and conversely.*

This has been proved by VERMES ([7], Theorem 3. III).

LEMMA 4. *If* μ_n *is a moment constant given by* (3), *then the sequence-to-sequence transformation* (H_2^*, μ) *is regular if, and only if,*

$$\int_0^1 \left|\frac{d\chi(u)}{u}\right| < \infty, \quad \text{and} \quad \int_0^1 \frac{d\chi(u)}{u} = 1.$$

Lemma 4 is given by HARDY ([1], Theorem 219) with the additional assumption of regularity of μ_n. This additional assumption is seen to be superfluous in Hardy's proof of Lemma 4 and it is in fact not fulfilled in the case of the circle method regarded as a particular case of (H_2^*, μ) summability ([1], § 14.21 (1)).

3. THEOREMS.

THEOREM 1. *Let $\sum_0^\infty a_n$ be any series belonging to the class $\mathfrak{B}$ of series with bounded partial sums. If $\sum a_n$ is summable to l by Abel's method, then the transform $\sum b_n$ of the series given by any quasi-Hausdorff α_0-matrix is also summable to l by Abel's method.*

PROOF. In our present transformation (H_1^*, μ), $\lambda_{n\,k}^*$ is defined by (1) and μ_n by (3) and (4), since the matrix of transformation is assumed to be an α_0-matrix and so satisfies the conditions of Lemma 2. Now,

$$b_n = \sum_\varkappa \lambda_{n\,\varkappa}^* a_\varkappa$$

and the matrix $\|\lambda^*\|$ is an α_0-matrix. Therefore by Lemma 3, there exists a T-matrix B corresponding to this and defined by (5). It is easily seen, as in the course of the proof of Theorem II of [4], that the T-matrix B corresponding to $\|\lambda^*\|$ is given by the equation

$$b_{n\varkappa} = \binom{\varkappa}{n} \Delta^{\varkappa-n} \mu_{n+1}.$$

Therefore we have now that if $\{s_n\}$ and $\{t_n\}$ are the sequences of partial sums of $\sum a_n$ and $\sum b_n$ respectively,

$$t_n = \sum_{\varkappa \geq n} \binom{\varkappa}{n} \Delta^{\varkappa-n} \mu_{n+1} s_\varkappa. \tag{6}$$

But, since μ_n is given by (3),

$$t_n = \sum_{\varkappa \geq n} \int_0^1 s_\varkappa \binom{\varkappa}{n} (1-u)^{\varkappa-n} u^{n+1} \, d\chi(u).$$

We have, consequently

$$\sum_n t_n x^n = \sum_n x^n \sum_{\varkappa \geq n} \int_0^1 s_\varkappa \binom{\varkappa}{n} (1-u)^{\varkappa-n} u^{n+1} \cdot d\chi(u) \tag{7}$$

provided the right-hand member exists. To prove this existence, consider the right-hand member with $s_\varkappa$ replaced by $|s_\varkappa|$ and $\chi(u)$, supposed to be monotonic increasing (as is permissible). The right-hand member with these changes, is

$$\begin{aligned}
&\sum_n x^n \sum_{\varkappa \geq n} \int_0^1 |s_\varkappa| \cdot \binom{\varkappa}{n} (1-u)^{\varkappa-n} u^{n+1} \, d\chi(u) \\
&= \sum_n x^n \int_0^1 \sum_{\varkappa \geq n} |s_\varkappa| \cdot \binom{\varkappa}{n} (1-u)^{\varkappa-n} u^{n+1} \, d\chi(u) \\
&= \int_0^1 \sum_n x^n \sum_{\varkappa \geq n} |s_\varkappa| \binom{\varkappa}{n} (1-u)^{\varkappa-n} u^{n+1} \, d\chi(u) \\
&= \int_0^1 \sum_{\varkappa=0}^\infty |s_\varkappa| \sum_{n=0}^k \binom{\varkappa}{n} (1-u)^{\varkappa-n} \cdot x^n \cdot u^{n+1} \cdot d\chi(u) \\
&= \int_0^1 \sum_\varkappa |s_\varkappa| (1-u+xu)^\varkappa \cdot u \cdot d\chi(u)
\end{aligned}$$

every inversion of operation being justified by the fact that we have only positive integrands or terms. The last integral is finite on account of the boundedness of $\{s_n\}$ and since the function $\chi(u)$ satisfies the conditions of Lemma 2.

Hence, we get from (7)

$$\sum t_n x^n = \int_0^1 \sum_\varkappa s_\varkappa (1-u+xu)^\varkappa \cdot u \cdot d\chi(u). \tag{8}$$

Put now

$$f(x) = \sum_n s_n x^n \quad \text{and} \quad F(x) = \sum_n t_n x^n.$$

Substituting $x = 1 - \frac{1}{y}$, we have

$$g(y) = f\left(1-\frac{1}{y}\right) = f(x) \quad \text{and} \quad G(y) = F\left(1-\frac{1}{y}\right) = F(x)$$

and

$$(1-x) f(x) \to l \quad \text{as} \quad x \to 1-0$$

or

$$\frac{1}{y} \cdot g(y) \to l \quad \text{as} \quad y \to \infty.$$

Also from equation (8) above, we have,

$$\begin{aligned} G(y) &= \int_0^1 f(1-u+xu) \cdot u \cdot d\chi(u) \\ &= \int_0^1 u \cdot f\left(1-\frac{u}{y}\right) d\chi(u) \\ &= \int_0^1 u \cdot g\left(\frac{y}{u}\right) d\chi(u) \end{aligned}$$

and

$$\frac{G(y)}{y} = \int_0^1 \frac{u}{y} \cdot g\left(\frac{y}{u}\right) d\chi(u).$$

Since $\frac{g(y)}{y} \to l$ as $y \to \infty$, the same is true of $\frac{g\left(\frac{y}{u}\right)}{\frac{y}{u}}$ also, since, for $0 \leq u \leq 1$, $\frac{y}{u} \geq y$.

Therefore, by Lebesgue's theorem on dominated convergence, $\frac{G(y)}{y} \to l$ as $y \to \infty$, since, by Lemma 2, $\chi(u)$ satisfies the conditions (4), i.e. $\sum b_n$ is Abel summable to l and this completes the proof of the theorem.

REMARK 1. In the special case when

$$\chi(u) = 0, \quad 0 \leq u < a, \quad \chi(u) = 1, \quad a \leq u \leq 1, \qquad \text{so that } \mu_n = a^n$$

which will give the circle method of summation, regarded as a series-to-series transformation, the right-hand members of (7) and (8) become respectively,

$$\sum_n x^n \sum_{\varkappa \geq n} \binom{\varkappa}{n} (1-a)^{\varkappa - n} a^{n+1} s_\varkappa \tag{7'}$$

and

$$(8)' \qquad \sum_{\varkappa} s_\varkappa (1-a+xa)^\varkappa \cdot a.$$

Now, on account of the absolute convergence of the series in (8)', ensured by the Abel summability of $\{s_n\}$, we can effect the interchange of order of summation in the steps between (7) and (8) and establish the equality of (7)' and (8)' irrespective of whether $\{s_n\}$ is bounded or not. Thus in the special case of the circle method $\sum_n t_n x^n$ exists as the series in (8)' or alternatively as the integral in (8). Hence the proof of Szász's theorem is contained in that of Theorem 1.

REMARK 2. The need for the restriction $s_n = O(1)$ becomes evident when we consider the series $\sum_n a_n$, $a_n \equiv 1$ and its transformation $\sum b_n$, where

$$b_n = \frac{a_n}{n+1} + \frac{a_{n+1}}{n+2} + \cdots$$

which does not exist.

Even the slightly weaker assumption that the radius of convergence of $\sum a_n x^n$ is unity will not suffice for the existence of $\{b_n\}$ is evident from the above example.

REMARK 3. Another restriction in Theorem 1 is that the transformation is by an α_0-matrix, which is stronger than assuming that it is by an α-matrix. It might however be noted that, in case it is not by an α_0-matrix, the existence of an equivalent T-matrix is not ensured, and therefore we cannot proceed to equation (6) and thence to (7). In fact we have then, instead of the right-hand member of (7) the double series

$$\sum_n x^n \sum_{\varkappa \geq n} a_\varkappa \int_0^1 \binom{\varkappa}{n} (1-u)^{\varkappa - n} u^n \, d\chi(u)$$

which need not be absolutely convergent, even if $s_n = O(1)$ as seen from the example $a_n = (-1)^n$ and $\chi(u) = u$, kindly supplied to me by the referee.

THEOREM 2. *Let s_n $(n = 0, 1, 2, \ldots)$ be a bounded sequence summable to l by Abel's method. Then a regular transform t_n $(n = 0, 1, 2, \ldots)$ of s_n by the method (H_2^*, μ) with moment constants μ_n, is also summable to l by Abel's method.*

PROOF. We proceed as in the proof of Theorem 1 and obtain instead of (8)

$$(1-x)\sum_n t_n x^n = (1-x)\int_0^1 \sum_\varkappa s_\varkappa (1-u+xu)^\varkappa \cdot d\chi(u)$$

where the series on the left is convergent (absolutely) for $0 < x \leq 1$ the series on the right, under the integral sign, being so. Appealing now to Lemma 4, exactly as we did formerly to Lemma 2, we see that the conclusion

$$\lim_{x \to 1-0} (1-x)\sum t_n x^n = l$$

follows from the hypothesis

$$\lim_{x \to 1-0} (1-x)\sum s_n x^n = l.$$

A special case of the transformation (H_2^*, μ) is the following transformation

$$C_1^* \quad t_n = (n+1)\left[\frac{s_n}{(n+1)(n+2)} + \frac{s_{n+1}}{(n+2)(n+3)} + \cdots\right]$$

where $\mu_n = (n+2)^{-1}$ and so $\chi(u) = \frac{u^2}{2}$, known to be equivalent to the $(C, 1)$ transformation of a sequence ([1], § 11.21 (3)).

In conclusion, I wish to thank Professors C. T. RAJAGOPAL and V. GANAPATHY IYER for their helpful guidance.

References

[1] HARDY, G. H.: Divergent Series. Oxford 1949. — [2] PATI, T.: Products of Summability Methods. Proc. Nat. Inst. Sci. Ind. **20**, 348—351 (1954). — [3] RAJAGOPAL, C. T.: Theorems on the Product of Two Summability Methods with Applications. J. Ind. Math. Soc. (New Series) **18**, 89—105 (1954). — [4] RAMANUJAN, M. S.: Series-to-series Quasi-Hausdorff Transformations. J. Ind. Math. Soc. (New Series) **17**, 47—53 (1953). — [5] SZASZ, O.: On Products of Summability Methods. Proc. Amer. Math. Soc. **3**, 257—263 (1952). — [6] SZASZ, O.: On the Product of Two Summability Methods. Ann. Soc. Polon. Math. **25**, 75—84 (1952/53). — [7] VERMES, P.: Conservative Series-to-series Transformation Matrices. Acta Sci. Math. Szeged. **15**, 23—38 (1951).

Ramanujan Institute of Mathematics, Madras (India)

(Eingegangen am 25. Februar 1956)

BAUER, H.
Math. Zeitschr. Bd. 65, S. 448—482 (1956)

Über die Beziehungen einer abstrakten Theorie des Riemann-Integrals zur Theorie Radonscher Maße*)

Von
HEINZ BAUER

Einleitung

Neben die abstrakte Integrationstheorie von STONE [*21*], welche die klassische Lebesguesche Theorie zum Vorbild hat, ist neuerdings eine abstrakte Theorie des Riemann-Integrals von LOOMIS [*14*] getreten. „Abstrakt" heißt hierbei, daß auf eine Topologie im Definitionsbereich der zu integrierenden Funktionen kein Bezug genommen wird. Ausgangspunkt der Untersuchungen von Loomis ist vielmehr, ähnlich wie bei Stone, ein positives, lineares Funktional I über einem System $\mathfrak{R}$ reeller Funktionen, deren gemeinsamer Definitionsbereich eine Menge E ist; hierbei wird $\mathfrak{R}$ als Vektorverband angenommen, welcher als Ersatz für die eventuell fehlende Konstante 1 mit jeder Funktion g auch $\min(1, g)$ enthält. Im Gegensatz zur Stoneschen Theorie wird die Stetigkeit des Funktionals I nicht gefordert. Ziel ist, wie bei allen Untersuchungen dieser Art, durch Einführung einer geeigneten additiven Mengenfunktion $\mu \mid \mathfrak{Q}$ und Definition eines μ-Unterteilungsintegrals eine Darstellung von $I(f)$ als μ-Integral für eine möglichst umfassende Klasse von Funktionen $f \in \mathfrak{R}$ zu finden. Die fehlende Stetigkeit von I führt hierbei zu einem neuen Gesichtspunkt; es wird nämlich ein gewisses Prinzip der Sparsamkeit bei der Bestimmung der μ-meßbaren Mengen verfolgt. Bei Loomis' ist $\mathfrak{Q}$ das System aller Teilmengen Q von E, deren charakteristische Funktionen von oben und unten her durch Funktionen aus $\mathfrak{R}$ beliebig genau I-einschließbar sind; die charakteristische Funktion χ_Q einer Menge $Q \in \mathfrak{Q}$ erhält dann bei jeder möglichen Fortsetzung von I zu einem positiven, linearen Funktional über einem χ_Q als Element enthaltenden Vektorraum von reellen Funktionen auf E den gleichen Funktionalwert, nämlich $\mu(Q)$ zugeordnet. Mit Hilfe von $\mu \mid \mathfrak{Q}$ gewinnt Loomis das Integral $\int f \, d\mu$ durch Bildung von Ober- und Untersummen nach dem Vorbild des klassischen Riemann-Integrals.

Wendet man die Loomissche Theorie speziell auf den „konkreten" Fall eines lokal kompakten Raumes E an und wählt man für $\mathfrak{R}$ die Menge $\mathfrak{K}(E)$ der auf E stetigen Funktionen mit kompaktem Träger, so ist I ein Radonsches Maß ϱ auf E. Es zeigt sich, daß in diesem Falle $\mathfrak{Q}$ das System der ϱ-quadrierbaren Mengen ist, d.h. der relativ kompakten Teilmengen von E, deren Begrenzung eine ϱ-Nullmenge ist. Ferner erweist sich das System der bezüglich I im Sinne von Loomis Riemann-integrierbaren Funktionen als das System aller ϱ-fast überall stetigen Funktionen auf E mit kompaktem

*) Von der Naturwissenschaftlichen Fakultät der Universität Erlangen als Habilitationsschrift angenommen.

Träger. Man gewinnt auf diese Weise aus der Theorie von Loomis die von HAUPT und PAUC [10], [19], [12b], S. 174ff., entwickelte Theorie des Riemann-Integrals in lokal kompakten Räumen.

Das Hauptergebnis der vorliegenden Arbeit ist gewissermaßen die Umkehrung dieser Feststellung. Zunächst läßt sich zeigen, daß es für die Loomissche Theorie keine Beschränkung der Allgemeinheit bedeutet, die Funktionen aus $\mathfrak{R}$ als beschränkt vorauszusetzen und anzunehmen, daß je zwei verschiedene Elemente von E durch eine Funktion aus $\mathfrak{R}$ getrennt werden können und in jedem Element von E mindestens eine Funktion aus $\mathfrak{R}$ nicht verschwindet. Unter diesen zusätzlichen Annahmen über E und $\mathfrak{R}$ wird dann die Existenz eines durch E und $\mathfrak{R}$ im wesentlichen eindeutig bestimmten, lokal kompakten Raumes $E'_{\mathfrak{R}}$ nachgewiesen, der E als dichte Teilmenge enthält und folgende Eigenschaft besitzt: Zu jedem positiven, linearen Funktional $I|\mathfrak{R}$ existiert ein eindeutig bestimmtes, positives Radonsches Maß ϱ' auf $E'_{\mathfrak{R}}$ derart, daß das Mengensystem $\mathfrak{Q}$ genau die Spur des Systems der ϱ'-quadrierbaren Teilmengen von $E'_{\mathfrak{R}}$ und ferner das System der bezüglich I Riemann-integrierbaren Funktionen $f|E$ gleich dem System der auf E verengerten, ϱ'-fast überall stetigen Funktionen $f'|E'_{\mathfrak{R}}$ mit kompaktem Träger ist. Für jede ϱ'-quadrierbare Menge Q' und deren Spur Q in E ist $\mu(Q)=\varrho'(Q')$; für jede ϱ'-fast überall stetige Funktion $f'|E'_{\mathfrak{R}}$ mit kompaktem Träger und deren Verengerung $f|E$ ist $\int f\,d\mu=\int f'\,d\varrho'$.

Durch diesen Darstellungssatz für die nach Loomis Riemann-integrierbaren Funktionen wird das nicht notwendig stetige Funktional I mit dem Radonschen Maß ϱ', also einem stetigen Funktional über dem Vektorraum $\mathfrak{R}(E'_{\mathfrak{R}})$ der auf $E'_{\mathfrak{R}}$ stetigen Funktionen mit kompaktem Träger in Beziehung gesetzt. Der Klärung des hier vorliegenden Sachverhalts dient eine Bemessung des Stetigkeitsdefektes von I durch die Zerlegung von I in seinen sog. stetigen Teil I_c und rein-unstetigen Teil I_p. Da aus der Stoneschen Theorie die Existenz einer Integraldarstellung für den stetigen Teil I_c folgt, lassen sich mit Hilfe dieser Zerlegung Bedingungen herleiten, unter welchen für eine vorgegebene Funktion $f\in\mathfrak{R}$ der Funktionalwert $I(f)$ als Integral darstellbar ist. Hierdurch erscheinen bekannte Sätze über Integraldarstellungen in neuem Lichte.

Die Arbeit gliedert sich in vier Paragraphen. § 1 gibt einen kurzen Überblick über die Theorie von Loomis und ergänzt sie durch einige auf unser Ziel ausgerichtete Bemerkungen. § 2 enthält einen Beweis für die Existenz der Zerlegung $I=I_c+I_p$ und bringt diese in Zusammenhang mit den Begriffsbildungen der Loomisschen Theorie. Insbesondere wird gezeigt, daß der Zerlegung von I die Zerlegung von μ in σ-additiven und rein-endlich-additiven Teil entspricht. In § 3 wird die Loomissche Theorie auf den konkreten Fall eines lokal kompakten Raumes angewandt als Vorstufe für den Beweis des Darstellungssatzes, der dann in § 4 erbracht wird. In einem Beispiel am Schluß der Arbeit wird die Stonesche Kompaktifizierung einer Booleschen Mengenalgebra als Raum $E'_{\mathfrak{R}}$ gedeutet[1]).

[1]) Eine Vorankündigung der wichtigsten Resultate dieser Arbeit (ohne Beweise) findet sich in [7].

§ 1. Riemann-Integrierbarkeit bezüglich eines linearen Funktionals. Abstrakter Fall[2])

1.1. Es sei E eine Menge beliebiger Elemente $x, y, \ldots$ und $\mathfrak{R}$ eine Menge reeller, endlicher, auf E definierter Funktionen mit folgenden Eigenschaften:

(1) *sind* $g, h \in \mathfrak{R}$ *und* α, β *reelle Zahlen, so ist* $\alpha g + \beta h \in \mathfrak{R}$;

(2) *ist* $g \in \mathfrak{R}$, *so ist* $|g| \in \mathfrak{R}$;

(3) *ist* $g \in \mathfrak{R}$, *so ist* $\min(1, g) \in \mathfrak{R}$[3]).

Zufolge (1) und (2) ist $\mathfrak{R}$ ein Vektorverband reeller Funktionen über dem Körper P der reellen Zahlen; insbesondere liegen mit je endlich vielen Funktionen $g_1, \ldots, g_n$ auch die Funktionen $\max(g_1, \ldots, g_n)$ und $\min(g_1, \ldots, g_n)$ in $\mathfrak{R}$. Zufolge (1) bis (3) liegt $\min(\alpha, g)$ in $\mathfrak{R}$ für jede Funktion $g \in \mathfrak{R}$ und jede reelle Zahl $\alpha \geq 0$.

An anderer Stelle [*17*] haben NÖBELING und Verf. kennzeichnende Eigenschaften derjenigen beschränkten, reellen Funktionen auf E angegeben, welche durch Funktionen aus $\mathfrak{R}$ gleichmäßig approximiert werden können. Wir entnehmen dieser Arbeit das folgende für uns wichtige Resultat[4]).

Lemma 1. *Die Menge* $\mathfrak{A}$ *aller beschränkten, reellen, auf* E *definierten Funktionen, die durch Funktionen aus* $\mathfrak{R}$ *gleichmäßig approximiert werden können, ist eine Algebra*[5]). *Zu jeder Funktion* $f \in \mathfrak{A}$ *existiert eine gegen* f *auf* E *gleichmäßig konvergente Folge* $\{g_n\}_{n=1,2,\ldots}$ *von (beschränkten) Funktionen aus* $\mathfrak{R}$ *mit folgender Eigenschaft:*

$$(4)\quad \begin{cases} 0 \leq g_1(x) \leq g_2(x) \leq \cdots \leq f(x) & \text{für jedes } x \in E \text{ mit } f(x) \geq 0; \\ 0 \geq g_1(x) \geq g_2(x) \geq \cdots \geq f(x) & \text{für jedes } x \in E \text{ mit } f(x) \leq 0. \end{cases}$$

1.2. Neben E und $\mathfrak{R}$ sei gegeben ein auf $\mathfrak{R}$ definiertes, reelles, *positives*, lineares Funktional I.

Als (zweiseitige) *Vervollständigung* $\overline{\mathfrak{R}} = \mathfrak{V}(\mathfrak{R}, I)$ von $\mathfrak{R}$ bezüglich I bezeichnen wir die Menge aller reellen Funktionen $f|E$, zu welchen bei beliebig vorgegebenem $\varepsilon > 0$ Funktionen $g, h \in \mathfrak{R}$ existieren mit

$$g \leq f \leq h \quad \text{und} \quad I(h - g) \leq \varepsilon\ ^{6}).$$

Dann besitzt $\overline{\mathfrak{R}}$ die Eigenschaften (1) bis (3) von $\mathfrak{R}$.

---

[2]) Die in den Nummern 1.2 bis 1.6 dieses Paragraphen erwähnten Begriffsbildungen finden sich bei LOOMIS [*14*].

[3]) Ist $F(u_1, \ldots, u_n)$ eine reelle Funktion in n reellen Variablen $u_1, \ldots, u_n$ und sind $g_1, \ldots, g_n$ auf E definierte, reelle Funktionen, so bezeichnet $F(g_1, \ldots, g_n)$ die durch die Zuordnung $x \to F(g_1(x), \ldots, g_n(x))$ auf E definierte Funktion. Es gilt also z. B. $|g|(x) = |g(x)|$ für alle $x \in E$.

[4]) Da sich die Eigenschaften (1) bis (3) von $\mathfrak{R}$ auf die Menge $\mathfrak{R}^*$ aller beschränkten Funktionen aus $\mathfrak{R}$ übertragen, folgt dies aus der Bemerkung zum Korollar des Kriteriums 2 und dem Satz 1 der zitierten Arbeit [*17*], S. 58 und 59.

[5]) Für je zwei Funktionen $f_1, f_2 \in \mathfrak{A}$ liegt also jede Linearkombination $\alpha f_1 + \beta f_2$ ($\alpha, \beta \in \mathsf{P}$) und das Produkt $f_1 \cdot f_2$ in $\mathfrak{A}$. Darüber hinaus besitzt $\mathfrak{A}$ auch die Eigenschaften (2) und (3) von $\mathfrak{R}$.

[6]) Wir schreiben $g \leq f \leq h$ für „$g(x) \leq f(x) \leq h(x)$ für alle $x \in E$".

Vermöge der Definition von $\overline{\mathfrak{R}}$ kann das Funktional I *auf genau eine Weise* zu einem positiven, linearen Funktional $\bar{I}|\overline{\mathfrak{R}}$ fortgesetzt werden. Für ein beliebiges $f \in \overline{\mathfrak{R}}$ gilt

$$\bar{I}(f) = \sup\big(I(g);\, g \leq f,\, g \in \mathfrak{R}\big) = \inf\big(I(h);\, h \geq f,\, h \in \mathfrak{R}\big). \tag{5}$$

$\overline{\mathfrak{R}}$ ist vollständig im Sinne der Gleichung $\mathfrak{V}(\overline{\mathfrak{R}}, \bar{I}) = \overline{\mathfrak{R}}$. Wir bezeichnen daher $\bar{I}|\overline{\mathfrak{R}}$ als die Vervollständigung des Funktionals $I|\mathfrak{R}$.

1.3. Für eine Teilmenge A von E bezeichne χ_A die charakteristische Funktion von A bezüglich der Grundmenge E. Eine Teilmenge Q von E heißt *quadrierbar* bezüglich I (kurz (I-)quadrierbar), wenn ihre charakteristische Funktion χ_Q ein Element von $\overline{\mathfrak{R}}$ ist. Das System aller I-quadrierbaren Mengen werde mit $\mathfrak{Q} = \mathfrak{Q}(\mathfrak{R}, I)$ bezeichnet; definitionsgemäß ist $\mathfrak{Q}(\overline{\mathfrak{R}}, \bar{I}) = \mathfrak{Q}(\mathfrak{R}, I)$. Aus den Eigenschaften (1) bis (3) von $\overline{\mathfrak{R}}$ (mit $\overline{\mathfrak{R}}$ an Stelle von $\mathfrak{R}$) folgt, daß $\mathfrak{Q}$ ein Boolescher Mengenverband ist, welcher im allgemeinen keine Einheit besitzt. Ist jedoch die auf E konstante Funktion $f = 1$ ein Element von $\mathfrak{R}$, so gilt $E \in \mathfrak{Q}$.

Die durch die Gleichung

$$\mu(Q) = \bar{I}(\chi_Q) \qquad (Q \in \mathfrak{Q}) \tag{6}$$

auf $\mathfrak{Q}$ definierte, reelle, endliche Mengenfunktion μ ist ein *vollständiger Inhalt*, d.h. $\mu|\mathfrak{Q}$ ist (endlich-)additiv und nicht negativ und es gilt $M \in \mathfrak{Q}$ für eine Menge $M \subseteq E$, wenn zu beliebig vorgegebenem $\varepsilon > 0$ Mengen $P, Q \in \mathfrak{Q}$ mit $P \subseteq M \subseteq Q$ und $\mu(Q - P) \leq \varepsilon$ existieren.

1.4. Mit $\mathfrak{E}$ bezeichnen wir die Menge aller endlichen Linearkombinationen von charakteristischen Funktionen I-quadrierbarer Mengen. Es besteht also $\mathfrak{E}$ aus allen reellen Funktionen der Gestalt $f = \sum_{\nu=1}^{n} \lambda_\nu \chi_{Q_\nu}$ $(Q_\nu \in \mathfrak{Q},\ \lambda_\nu \in \mathsf{P})$. Auch $\mathfrak{E}$ besitzt die Eigenschaften (1) bis (3) von $\mathfrak{R}$ und ist überdies eine Algebra.

Wegen $\mathfrak{E} \subseteq \overline{\mathfrak{R}}$ ist das Funktional $\bar{I}|\mathfrak{E}$ als Verengerung von $\bar{I}|\overline{\mathfrak{R}}$ und folglich auch die Vervollständigung $\mathfrak{S} = \mathfrak{V}(\mathfrak{E}, \bar{I})$ von $\mathfrak{E}$ bezüglich $\bar{I}$ definiert. $\mathfrak{S}$ besitzt daher ebenfalls die Eigenschaften (1) bis (3) von $\mathfrak{R}$ und ist ebenso wie $\mathfrak{E}$ eine Algebra von Funktionen. Wegen $\mathfrak{E} \subseteq \mathfrak{S} \subseteq \overline{\mathfrak{R}}$ ist die Verengerung $\bar{I}|\mathfrak{S}$ von $\bar{I}|\overline{\mathfrak{R}}$ gleich der Vervollständigung des Funktionals $\bar{I}|\mathfrak{E}$.

Nach Loomis [*14*] heißen die Funktionen aus $\mathfrak{S}$ (eigentlich) *Riemann-integrierbar* bezüglich I; die einer Funktion $f \in \mathfrak{S}$ zugeordnete reelle Zahl $\bar{I}(f)$ heißt *Riemann-Integral* von f bezüglich I. Da $\bar{I}(f)$ durch Bildung von Ober- und Untersummen nach dem Vorbild des klassischen Riemann-Integrals definiert werden kann, setzt man

$$\bar{I}(f) = \int f\, d\mu \qquad (f \in \mathfrak{S}). \tag{7}$$

1.5. Eine nicht negative, endliche, reelle Funktion $f|E$ heißt (I-)*meßbar*, wenn die Urbildmenge

$$[f > \lambda] = \{x : f(x) > \lambda\}$$

I-quadrierbar ist für alle reellen Zahlen $\lambda > 0$, abgesehen von höchstens abzählbar unendlich vielen Ausnahmen. Eine beliebige, endliche, reelle Funktion $f|E$ heißt (I-)meßbar, wenn ihr Positivteil $f^+ = \max(f, 0)$ und ihr Negativteil $f^- = (-f)^+$ meßbar sind.

Da $\mathfrak{R}$ die Eigenschaft (3) besitzt gilt nach LOOMIS [14], S. 171:

Satz 1. *Die Funktionen aus $\overline{\mathfrak{R}}$ sind sämtlich I-meßbar. Insbesondere sind die Funktionen aus $\mathfrak{R}$ meßbar in bezug auf jedes auf $\mathfrak{R}$ definierte, positive, lineare Funktional.*

Zusatz. *Für eine nicht negative Funktion $f \in \overline{\mathfrak{R}}$ gilt auch $[f \geq \lambda] \in \mathfrak{Q}$ für alle $\lambda > 0$, abgesehen von höchstens abzählbar unendlich vielen Ausnahmewerten.*

Mit Hilfe des eingeführten Meßbarkeitsbegriffes können die Riemann-integrierbaren Funktionen wie folgt gekennzeichnet werden.

Satz 2. *Notwendig und hinreichend für die Riemann-Integrierbarkeit einer reellen Funktion $f|E$ bezüglich des Funktionals I ist folgende Bedingung: f ist beschränkt, I-meßbar und im Komplement einer I-quadrierbaren Menge gleich Null.*

Beweis. Es sei f Riemann-integrierbar. Wegen $\mathfrak{E} \subseteq \overline{\mathfrak{R}}$ ist $f \in \overline{\mathfrak{R}}$ und folglich nach Satz 1 meßbar. Die Funktionen aus $\mathfrak{E}$ sind als „$\mathfrak{Q}$-Treppenfunktionen" beschränkt und im Komplement einer quadrierbaren Menge gleich Null. Die gleichen Eigenschaften besitzt dann auch f als Element der Vervollständigung $\overline{\mathfrak{E}}$ von $\mathfrak{E}$ bezüglich $\bar{I}$. Also ist die Bedingung erfüllt. Nach LOOMIS [14], S. 172 folgt aus dem Bestehen der Bedingung umgekehrt die Riemann-Integrierbarkeit von f.

In diesem Zusammenhang beweisen wir das

Lemma 2. *Eine reelle Funktion $f|E$ ist dann und nur dann im Komplement einer quadrierbaren Menge gleich Null, wenn eine nicht negative Funktion $h \in \mathfrak{R}$ derart existiert, daß $f(x) = 0$ ist für jedes $x \in E$ mit $h(x) < 1$.*

Beweis. Die Funktion f sei im Komplement der quadrierbaren Menge Q gleich Null. Zu Q existiert eine Funktion $h \in \mathfrak{R}$ mit $\chi_Q \leq h$; diese leistet das Verlangte. — Umgekehrt existiere zu f eine nicht negative Funktion $h \in \mathfrak{R}$ mit der genannten Eigenschaft. Nach Satz 1 ist h meßbar. Daher existiert eine reelle Zahl λ mit $0 \leq \lambda < 1$, für welche die Urbildmenge $Q = [h > \lambda]$ quadrierbar ist. Für ein beliebiges nicht in Q gelegenes Element $x \in E$ ist $h(x) \leq \lambda < 1$ und folglich $f(x) = 0$. Also ist f im Komplement von Q gleich Null.

Mit Hilfe dieses Lemmas ergeben sich nun zwei Korollare zu Satz 2.

Korollar 1. *Notwendig und hinreichend dafür, daß die beschränkten Funktionen aus $\mathfrak{R}$ sämtlich bezüglich I und damit bezüglich eines jeden auf $\mathfrak{R}$ definierten, positiven, linearen Funktionals Riemann-integrierbar sind, ist folgende Bedingung:*

(8) *Zu jeder Funktion $g \in \mathfrak{R}$ existiert eine nicht negative Funktion $h \in \mathfrak{R}$ derart, daß $g(x) = 0$ ist für jedes $x \in E$ mit $h(x) < 1$*[7]).

[7]) Die Bedingung (8) tritt in anderem Zusammenhang auch bei ARENS [1] auf.

Beweis. Wegen Satz 2 und Lemma 2 ist nur noch zu zeigen, daß die Bedingung (8) notwendig ist. Hierzu sei g eine Funktion aus $\mathfrak{R}$. Wir setzen $g' = \min(1, g^+) - \min(1, g^-)$, wobei g^+ den Positiv- und g^- den Negativteil von g bezeichnet. Dann ist g' eine beschränkte Funktion aus $\mathfrak{R}$. Nach Satz 2 und Lemma 2 existiert also eine nicht negative Funktion $h \in \mathfrak{R}$ derart, daß aus $h(x) < 1$ stets $g'(x) = 0$ folgt. Die Funktion h leistet auch bezüglich g das in (8) Verlangte: Für ein $x \in E$ ist entweder $g^+(x) = 0$ oder $g^-(x) = 0$; folglich zieht $g'(x) = 0$ die Gleichungen $g^+(x) = g^-(x) = g(x) = 0$ nach sich.

Korollar 2. *Eine beschränkte, durch Funktionen aus $\mathfrak{R}$ gleichmäßig approximierbare, reelle Funktion $f|E$ ist dann und nur dann Riemann-integrierbar bezüglich I und damit bezüglich eines jeden auf $\mathfrak{R}$ definierten, positiven, linearen Funktionals, wenn eine nicht negative Funktion $h \in \mathfrak{R}$ existiert derart, daß $f(x) = 0$ ist für jedes $x \in E$ mit $h(x) < 1$.*

Beweis. Wegen der Sätze 1 und 2 sowie wegen Lemma 2 genügt es zu zeigen, daß die Funktion f in $\overline{\mathfrak{R}}$ liegt, falls zu f ein $h \in \mathfrak{R}$ mit den genannten Eigenschaften existiert. Hierzu kann ohne Beschränkung der Allgemeinheit $f \geqq 0$ angenommen werden, da mit f auch f^+ und f^- beschränkt und durch Funktionen aus $\mathfrak{R}$ gleichmäßig approximierbar sind und da aus $h(x) < 1$ auch $f^+(x) = f^-(x) = 0$ folgt. Zu $f \geqq 0$ und einer beliebigen reellen Zahl $\varepsilon > 0$ existiert aber nach Lemma 1 eine Funktion $g \in \mathfrak{R}$ mit $0 \leqq g \leqq f \leqq g + \varepsilon$. Da aus $f(x) \neq 0$ für ein $x \in E$ folgt $h(x) \geqq 1$, gilt dann sogar $g' \leqq f \leqq h'$, wobei $g' = g$ und $h' = g + \varepsilon h$ Funktionen aus $\mathfrak{R}$ sind. Wegen $I(h' - g') = \varepsilon I(h)$ liegt also f in $\overline{\mathfrak{R}}$.

1.6. Eine nicht negative, endliche, reelle Funktion $f|E$ heißt *uneigentlich Riemann-integrierbar* (bezüglich des Funktionals I), wenn f I-meßbar und die Menge der für alle $g \in \mathfrak{S}$ mit $g \leqq f$ gebildeten Zahlen $\int g\, d\mu$ nach oben beschränkt ist. Als uneigentliches Riemann-Integral $\overset{(u)}{\int} f\, d\mu$ von f wird die reelle Zahl

$$\overset{(u)}{\int} f\, d\mu = \sup\left(\int g\, d\mu;\ g \leqq f,\ g \in \mathfrak{S}\right) \tag{9}$$

definiert. Eine beliebige reelle Funktion $f|E$ heißt uneigentlich Riemann-integrierbar, wenn ihr Positiv- und Negativteil uneigentlich Riemann-integrierbar sind. Man setzt

$$\overset{(u)}{\int} f\, d\mu = \overset{(u)}{\int} f^+\, d\mu - \overset{(u)}{\int} f^-\, d\mu. \tag{10}$$

Aus Satz 2 folgt, daß die Menge $\mathfrak{S}$ der eigentlich Riemann-integrierbaren Funktionen ein *gesättigter*, linearer Unterraum von $\overline{\mathfrak{R}}$ ist, d.h. ein linearer Unterraum mit folgender Eigenschaft:

(11) *Aus $|g| \leqq |f|$, wobei $f \in \mathfrak{S}$ und $g \in \overline{\mathfrak{R}}$ ist, folgt $g \in \mathfrak{S}$.*

Also sind die Begriffsbildungen von [5] anwendbar. Danach heißt ein auf $\overline{\mathfrak{R}}$ definiertes, positives, lineares Funktional L *$\mathfrak{S}$-singulär*, wenn

$$L(f) = \sup\left(L(g);\ g \leqq f,\ g \in \mathfrak{S}\right) \tag{12}$$

für jede nicht negative Funktion $f \in \overline{\mathfrak{R}}$ gilt.

Satz 3. *Die Funktionen aus $\mathfrak{R}$ sind sämtlich uneigentlich Riemann-integrierbar bezüglich I. Das somit für alle $f\in\mathfrak{R}$ definierte uneigentliche Riemann-Integral $\overset{(u)}{\int} f\,d\mu$ ist die einzig mögliche Fortsetzung des auf $\mathfrak{S}$ definierten, positiven, linearen Funktionals $g\to\int g\,d\mu$ zu einem $\mathfrak{S}$-singulären, positiven, linearen Funktional auf $\mathfrak{R}$. Hierbei gilt für jedes $f\in\mathfrak{R}$*

$$\left|\overset{(u)}{\int} f\,d\mu\right|\leq I(|f|). \tag{13}$$

Beweis. Es sei f eine nicht negative Funktion aus $\mathfrak{R}$. Zufolge (7) gilt dann $\int g\,d\mu\leq I(f)$ für jedes $g\in\mathfrak{S}$ mit $g\leq f$. Da f nach Satz 1 auch meßbar ist, so ist also f uneigentlich Riemann-integrierbar. Gleiches gilt dann für eine beliebige Funktion $f\in\mathfrak{R}$, da mit f auch f^+ und f^- Elemente von $\mathfrak{R}$ sind. Für nicht negative Funktionen $f', f''\in\mathfrak{R}$ folgt aus (9) $\overset{(u)}{\int}(f'+f'')\,d\mu=\overset{(u)}{\int} f'\,d\mu+\overset{(u)}{\int} f''\,d\mu$. Hierzu genügt es zu bemerken, daß zu jeder Funktion $g\in\mathfrak{S}$ mit $0\leq g\leq f'+f''$ Funktionen $g', g''\in\mathfrak{S}$ mit $g=g'+g''$ sowie $0\leq g'\leq f'$ und $0\leq g''\leq f''$ existieren; wegen (11) leisten z.B. $g'=\min(g, f')$ und $g''=\max(0, g-f')$ das Verlangte. Aus der Gültigkeit der Gleichung

$$\overset{(u)}{\int}(f'+f'')\,d\mu=\overset{(u)}{\int} f'\,d\mu+\overset{(u)}{\int} f''\,d\mu$$

und $\overset{(u)}{\int} f'\,d\mu\geq 0$ für beliebige $f', f''\in\mathfrak{R}$ mit $f'\geq 0$ und $f''\geq 0$ sowie aus (10) folgt aber (vgl. [8], S. 34), daß $f\to\overset{(u)}{\int} f\,d\mu$ auf $\mathfrak{R}$ ein positives, lineares Funktional ist. Daß dieses die einzig mögliche $\mathfrak{S}$-singuläre Fortsetzung des auf $\mathfrak{S}$ erklärten, eigentlichen Riemann-Integrals ist, folgt aus (9) und (10). Aus (9) folgt (13) zunächst für den Spezialfall einer Funktion $f\geq 0$ aus $\mathfrak{R}$ und hieraus allgemein bei Beachtung von (10).

§ 2. Zerlegungssätze

2.1. Es sei $\mathfrak{M}$ ein Vektorverband von auf einer Menge A definierten, endlichen, reellen Funktionen und L ein auf $\mathfrak{M}$ definiertes, positives, lineares Funktional. Dann heißt L bekanntlich *stetig*, wenn es der folgenden Bedingung genügt: Für jede nicht fallende Folge $\{f_n\}_{n=1,2,\ldots}$ von Funktionen aus $\mathfrak{M}$, deren obere Einhüllende $f=\sup f_n$ ein Element von $\mathfrak{M}$ ist, gilt $L(f)=\sup L(f_n)$. Dagegen heißt L *rein-unstetig*, wenn $K=0$ das einzige stetige, (positive) lineare Funktional auf $\mathfrak{M}$ ist mit $0\leq K(f)\leq L(f)$ für jede Funktion $f\geq 0$ aus $\mathfrak{M}$.

Nach einem in [3], S. 116 von uns bewiesenen Zerlegungssatz existieren zu einem positiven, linearen Funktional $L|\mathfrak{M}$ genau ein stetiges, positives, lineares Funktional $L_c|\mathfrak{M}$ und genau ein rein-unstetiges, positives lineares Funktional $L_p|\mathfrak{M}$ derart, daß $L=L_c+L_p$ ist. Es heißt L_c der *stetige* und L_p der *rein-unstetige Teil* von L.

Drücken wir den Sachverhalt, daß $\{f_n\}_{n=1,2,\ldots}$ eine nicht fallende Folge reeller Funktionen auf A mit der oberen Einhüllenden f ist, durch das Zeichen $f_n\uparrow f$ aus und bezeichnen wir mit $\mathfrak{M}_+$ die Menge aller $f\in\mathfrak{M}$ mit $f\geq 0$, so gilt

Lemma 3. *Für jede Funktion* $f \in \mathfrak{M}_+$ *ist*

$$L_c(f) = \inf\left(\sup L(f_n);\ f_n \uparrow f,\ f_n \in \mathfrak{M}_+\right).$$

Beweis. Wir setzen $L^*(f) = \inf\left(\sup L(f_n);\ f_n \uparrow f,\ f_n \in \mathfrak{M}_+\right)$ für eine beliebige Funktion $f \in \mathfrak{M}_+$ und behaupten zunächst, daß L^* auf genau eine Weise zu einem positiven, linearen Funktional auf $\mathfrak{M}$ fortgesetzt werden kann, welches wir ebenfalls mit L^* bezeichnen. Hierfür ist notwendig und hinreichend, daß (a) $L^*(f+g) = L^*(f) + L^*(g)$ und (b) $L^*(f) \geq 0$ ist für beliebige Funktionen $f, g \in \mathfrak{M}_+$. Während (b) unmittelbar klar ist, folgt (a) aus der Bemerkung, daß die Existenz einer Folge von Funktionen $h_n \in \mathfrak{M}_+$ mit $h_n \uparrow f+g$ äquivalent ist mit der Existenz zweier Folgen von Funktionen $f_n \in \mathfrak{M}_+$ und $g_n \in \mathfrak{M}_+$ mit $h_n = f_n + g_n$ sowie $f_n \uparrow f$ und $g_n \uparrow g$. Bei gegebener Folge $\{h_n\}$ leisten nämlich $f_n = \min(h_n, f)$ und $g_n = h_n - f_n = \max(0, h_n - f)$ das Verlangte; aus $f_n \uparrow f$ und $g_n \uparrow g$ folgt umgekehrt $f_n + g_n \uparrow f+g$.

Wir behaupten weiter, daß $L^* | \mathfrak{M}$ stetig ist. Es genügt hierfür zu zeigen, daß $L^*(f) \leq \sup L^*(f_n)$ ist für jede Folge $\{f_n\}_{n=1,2,\ldots}$ von Funktionen aus $\mathfrak{M}_+$ mit $f_n \uparrow f$ und $f \in \mathfrak{M}_+$. Nach Definition von L^* existiert zu beliebigem $\varepsilon > 0$ und für jedes $n = 1, 2, \ldots$ eine Folge $\{g_{nk}\}_{k=1,2,\ldots}$ von Funktionen aus $\mathfrak{M}_+$ mit $g_{nk} \uparrow f_n$ und

$$\sup_{k=1,2,\ldots} L(g_{nk}) \leq L^*(f_n) + \frac{\varepsilon}{2^n}.$$

Wir setzen $f_{nk} = \max(g_{1k}, \ldots, g_{nk})$ und zeigen mittels vollständiger Induktion nach n die Gültigkeit der Ungleichung

$$\sup_{k=1,2,\ldots} L(f_{nk}) \leq L^*(f_n) + \varepsilon\left(1 - \frac{1}{2^n}\right) \qquad (n = 1, 2, \ldots).$$

Für $n = 1$ ist nichts zu zeigen. Der Schluß von n auf $n+1$ verläuft so: Setzt man $h_{nk} = \min(f_{nk}, g_{n+1,k})$, so ist $f_{n+1,k} = \max(f_{nk}, g_{n+1,k}) = f_{nk} + g_{n+1,k} - h_{nk}$ und folglich

$$L(f_{n+1,k}) = L(f_{nk}) + L(g_{n+1,k}) - L(h_{nk})$$

$$\leq L^*(f_n) + \varepsilon\left(1 - \frac{1}{2^n}\right) + L^*(f_{n+1}) + \frac{\varepsilon}{2^{n+1}} - L(h_{nk});$$

wegen $h_{nk} \uparrow f_n$ ist $L^*(f_n) \leq \sup_{k=1,2,\ldots} L(h_{nk})$ und daher, wie behauptet,

$$\sup_{k=1,2,\ldots} L(f_{n+1,k}) \leq L^*(f_{n+1}) + \varepsilon\left(1 - \frac{1}{2^{n+1}}\right).$$

Nach Konstruktion gilt $f_{nk} \leq f_{n+1,k+1}$ und entsprechend $L(f_{nk}) \leq L(f_{n+1,k+1})$ für beliebige $n, k = 1, 2, \ldots$. Es ist also $\{f_{nn}\}_{n=1,2,\ldots}$ eine nicht fallende Folge von Funktionen aus $\mathfrak{M}_+$ mit $\sup f_{nn} = \sup_n \left(\sup_k f_{nk}\right) = f$ und $\sup L(f_{nn}) = \sup_n \left(\sup_k L(f_{nk})\right)$[8]. Aus der für alle $n = 1, 2, \ldots$ bewiesenen Ungleichung folgt also nunmehr

$$L^*(f) \leq \sup L(f_{nn}) \leq \sup L^*(f_n) + \varepsilon$$

für jedes $\varepsilon > 0$ und hieraus $L^*(f) \leq \sup L^*(f_n)$. Also ist L^* stetig.

[8]) Vgl. Haupt-Aumann-Pauc [12a], S. 147, Nr. 5.4.4.1.

Nach Definition von L^* ist $L-L^*$ ein positives, lineares Funktional auf $\mathfrak{M}$. Wir zeigen, daß $L-L^*$ rein-unstetig ist. Hierzu sei $K|\mathfrak{M}$ ein stetiges, positives, lineares Funktional mit $0 \leq K(f) \leq L(f)-L^*(f)$ für alle $f \in \mathfrak{M}_+$. Für eine beliebige Folge $\{f_n\}_{n=1,2,\ldots}$ von Funktionen aus $\mathfrak{M}_+$ mit $f_n \uparrow f$ und $f \in \mathfrak{M}_+$ gilt dann die Ungleichung $0 \leq K(f) \leq \sup L(f_n) - L^*(f)$ wegen der Stetigkeit von K und L^*. Hieraus folgt $K(f)=0$ nach Definition von L^*. Also ist $L-L^*$ rein-unstetig.

Aus den bewiesenen Eigenschaften von L^* folgt $L^*=L_c$ und damit die Behauptung.

2.2. Wie in § 1 sei nun wieder ein System $\mathfrak{R}$ auf einer Menge E definierter, reeller Funktionen mit den Eigenschaften (1) bis (3) und ein positives, lineares Funktional $I|\mathfrak{R}$ vorgegeben.

Bezeichnet I_c den stetigen und I_p den rein-unstetigen Teil von $I|\mathfrak{R}$, so ergeben sich die zweiseitigen Vervollständigungen

$$\overline{\mathfrak{R}}_c = \mathfrak{V}(\mathfrak{R}, I_c) \quad \text{und} \quad \overline{\mathfrak{R}}_p = \mathfrak{V}(\mathfrak{R}, I_p)$$

von $\mathfrak{R}$ bezüglich I_c und I_p sowie die Booleschen Mengenverbände

$$\mathfrak{Q}_c = \mathfrak{Q}(\mathfrak{R}, I_c) \quad \text{und} \quad \mathfrak{Q}_p = \mathfrak{Q}(\mathfrak{R}, I_p)$$

bestehend aus allen I_c- bzw. I_p-quadrierbaren Teilmengen von E. Schließlich ergeben sich noch die Systeme $\mathfrak{S}_c$ und $\mathfrak{S}_p$ der bezüglich I_c bzw. I_p Riemann-integrierbaren Funktionen. Wir behaupten die Gültigkeit der folgenden drei Gleichungen[9]:

$$\overline{\mathfrak{R}} = \overline{\mathfrak{R}}_c \cap \overline{\mathfrak{R}}_p; \tag{14}$$

$$\mathfrak{Q} = \mathfrak{Q}_c \cap \mathfrak{Q}_p; \tag{15}$$

$$\mathfrak{S} = \mathfrak{S}_c \cap \mathfrak{S}_p. \tag{16}$$

Zum Beweis von (14) sei $f \in \overline{\mathfrak{R}}$. Zu beliebigem $\varepsilon > 0$ existieren dann Funktionen $g, h \in \mathfrak{R}$ mit $g \leq f \leq h$ und $I(h-g) \leq \varepsilon$. Wegen $I_c(h-g) \leq I(h-g)$ und $I_p(h-g) \leq I(h-g)$ ist also $f \in \overline{\mathfrak{R}}_c \cap \overline{\mathfrak{R}}_p$. Umgekehrt sei $f \in \overline{\mathfrak{R}}_c \cap \overline{\mathfrak{R}}_p$. Dann existieren zu beliebigem $\varepsilon > 0$ Funktionen $g', h', g'', h'' \in \mathfrak{R}$ mit $g' \leq f \leq h'$ und $g'' \leq f \leq h''$ derart, daß $I_c(h'-g') \leq \varepsilon$ und $I_p(h''-g'') \leq \varepsilon$ ist. Setzen wir also $g = \max(g', g'')$ und $h = \min(h', h'')$, so sind g und h Funktionen aus $\mathfrak{R}$ mit $g \leq f \leq h$ und $h-g \leq \min(h'-g', h''-g'')$. Folglich ist $I(h-g) = I_c(h-g) + I_p(h-g) \leq 2\varepsilon$, also $f \in \overline{\mathfrak{R}}$. — Aus (14) folgt (15), da $Q \in \mathfrak{Q}$ gleichbedeutend mit $\chi_Q \in \overline{\mathfrak{R}}$ ist. — Eine reelle Funktion $f|E$, die sowohl im Komplement einer I_c-quadrierbaren Menge als auch im Komplement einer I_p-quadrierbaren Menge gleich Null ist, ist nach Lemma 2 auch im Komplement einer I-quadrierbaren Menge gleich Null. Beachtet man dies, so folgt (16) aus (15) und Satz 2.

[9]) Bei deren Beweis wird nur benützt, daß I_c und I_p positive, lineare Funktionale auf $\mathfrak{R}$ mit $I = I_c + I_p$ sind.

2.3. Nach 1.2 können die Funktionale $I_c|\mathfrak{R}$ und $I_p|\mathfrak{R}$ zu positiven, linearen Funktionalen $\overline{I}_c|\overline{\mathfrak{R}}_c$ und $\overline{I}_p|\overline{\mathfrak{R}}_p$ vervollständigt werden; wegen (14) ist $\overline{\mathfrak{R}}$ der Durchschnitt der Definitionsbereiche von $\overline{I}_c$ und $\overline{I}_p$. Nach 1.3 leiten sich aus I_c und I_p die vollständigen Inhalte $\mu_c|\mathfrak{Q}_c$ und $\mu_p|\mathfrak{Q}_p$ ab, welche wegen (15) insbesondere auf $\mathfrak{Q}$ definiert sind. Der folgende Satz enthält unter anderem die Aussage, daß $\mu_c|\mathfrak{Q}$ der σ-additive und $\mu_p|\mathfrak{Q}$ der rein-endlich-additive Teil von $\mu|\mathfrak{Q}$ ist.

Hierzu sei bemerkt, daß eine auf einem Booleschen Verband $\mathfrak{B}$ von Teilmengen einer Grundmenge A definierte, nicht negative, additive Mengenfunktion ϱ *σ-additiv* heißt, wenn $\sup \varrho(M_n) = \varrho(M)$ ist für jede nicht fallende Folge $\{M_n\}_{n=1,2,\ldots}$ von Mengen aus $\mathfrak{B}$ mit $M = \bigcup_{n=1}^{\infty} M_n \in \mathfrak{B}$. Dagegen heißt ϱ *rein-endlich-additiv*, wenn $\tau = 0$ die einzige σ-additive Funktion $\tau|\mathfrak{B}$ ist mit $0 \leqq \tau(M) \leqq \varrho(M)$ für alle $M \in \mathfrak{B}$. Nach Yosîda und Hewitt [23][10]) existieren zu jeder nicht negativen, additiven Funktion $\varrho|\mathfrak{B}$ genau eine nicht negative, σ-additive Funktion $\varrho_c|\mathfrak{B}$ und genau eine nicht negative, rein-endlich-additive Funktion $\varrho_p|\mathfrak{B}$ mit $\varrho = \varrho_c + \varrho_p$. Man nennt ϱ_c den *σ-additiven* und ϱ_p den *rein-endlich-additiven Teil* von ϱ. In Analogie zu der in Lemma 3 behaupteten Formel gilt nach Woodbury [22][11])

$$\varrho_c(M) = \inf\left(\sup \varrho(M_n);\ M_n \uparrow M,\ M_n \in \mathfrak{B}\right) \tag{17}$$

für alle $M \in \mathfrak{B}$. Hierbei steht $M_n \uparrow M$ für $M_1 \subseteqq M_2 \subseteqq \cdots \subseteqq M = \bigcup_{n=1}^{\infty} M_n$.

Der bereits angekündigte Satz lautet nun:

Satz 4. I. *Es ist $\overline{I}_c|\overline{\mathfrak{R}}$ der stetige und $\overline{I}_p|\overline{\mathfrak{R}}$ der rein-unstetige Teil des Funktionals $\bar{I}|\overline{\mathfrak{R}}$; also gilt auf $\overline{\mathfrak{R}}$*

$$\overline{I}_c = \bar{I}_c \quad \text{und} \quad \overline{I}_p = \bar{I}_p. \tag{18}$$

II. *Es ist $\mu_c|\mathfrak{Q}$ der σ-additive und $\mu_p|\mathfrak{Q}$ der rein-endlich-additive Teil von $\mu|\mathfrak{Q}$.*

Beweis. *Zu* I. Wir zeigen zunächst, daß das Funktional $\overline{I}_c|\overline{\mathfrak{R}}_c$ stetig ist; insbesondere ist dann $\overline{I}_c|\overline{\mathfrak{R}}$ stetig. Hierzu sei $f \in \overline{\mathfrak{R}}_c$ und $\{f_n\}_{n=1,2,\ldots}$ eine Folge von Funktionen aus $\overline{\mathfrak{R}}_c$ mit $f_n \uparrow f$; ferner sei $\varepsilon > 0$ eine reelle Zahl. Dann existiert ein $g \in \mathfrak{R}$ mit $g \leqq f$ und $\overline{I}_c(f) - \varepsilon \leqq I_c(g)$, sowie für $n = 1, 2, \ldots$ ein $h_n \in \mathfrak{R}$ mit $f_n \leqq h_n$ und $I_c(h_n) \leqq \overline{I}_c(f_n) + \varepsilon/2^n$. Setzen wir $g_n = \max(h_1, \ldots, h_n)$ für jedes n, so ist $\{g_n\}$ eine nicht fallende Folge von Funktionen aus $\mathfrak{R}$ mit $f_n \leqq g_n$ und $I_c(g_n) \leqq \overline{I}_c(f_n) + \varepsilon(1 - 1/2^n)$; diese Ungleichung beweist man wie die hierzu analoge Ungleichung im Beweis von Lemma 3 durch vollständige Induktion nach n. Nun ist $g \leqq f \leqq \sup g_n$ und folglich $I_c(g) \leqq \sup I_c(g_n)$ wegen der Stetigkeit von $I_c|\mathfrak{R}$. Hieraus folgt $\overline{I}_c(f) - \varepsilon \leqq I_c(g) \leqq \sup I_c(g_n) \leqq \sup \overline{I}_c(f_n) + \varepsilon$; also ist $\overline{I}_c(f) \leqq \sup \overline{I}_c(f_n)$. Wegen der Positivität von $\overline{I}_c|\overline{\mathfrak{R}}_c$ folgt hieraus die behauptete Stetigkeit.

[10]) Wie in [3], S. 113 von Verf. gezeigt wurde, ist die bei Yosida-Hewitt gemachte Annahme der Existenz der Einheit A in V unnötig.

[11]) Ein Beweis dieser Formel unter allgemeineren Voraussetzungen findet sich in [6], S. 219.

Das Funktional $\overline{I}_p|\overline{\mathfrak{R}}$ (und damit auch $\overline{I}_p|\overline{\mathfrak{R}}_p$) ist rein-unstetig. Es sei nämlich $K|\overline{\mathfrak{R}}$ ein stetiges, positives, lineares Funktional mit $0 \leq K(f) \leq \overline{I}_p(f)$ für alle $f \geq 0$ aus $\overline{\mathfrak{R}}$. Da $\overline{I}_p(g) = I_p(g)$ für alle $g \in \mathfrak{R}$ und $I_p|\mathfrak{R}$ rein-unstetig ist, folgt dann zunächst $K(g) = 0$ für alle $g \in \mathfrak{R}$ und hieraus $K(f) = 0$ für alle $f \in \overline{\mathfrak{R}}$. Man hat hierzu nur zu berücksichtigen, daß $\overline{\mathfrak{R}} \subseteq \mathfrak{B}(\mathfrak{R}, K)$ ist.

Schließlich ist $\overline{I}_c + \overline{I}_p$ eine Fortsetzung von $I|\mathfrak{R}$ zu einem positiven, linearen Funktional auf $\overline{\mathfrak{R}}$. Da $\overline{I}|\overline{\mathfrak{R}}$ nach 1.2 die einzige derartige Fortsetzung ist, gilt $\overline{I} = \overline{I}_c + \overline{I}_p$ auf $\overline{\mathfrak{R}}$. Also ist $\overline{I}_c = I_c$ und $\overline{I}_p = I_p$ auf $\overline{\mathfrak{R}}$, wie behauptet wurde.

Zu II. Es sei $\{Q_n\}_{n=1,2,\ldots}$ eine Folge von Mengen aus $\mathfrak{Q}_c$ mit $Q_n \uparrow Q$ und $Q \in \mathfrak{Q}_c$. Dann gilt $\chi_{Q_n} \uparrow \chi_Q$ für die Folge der zugehörigen charakteristischen Funktionen. Zufolge der bereits bewiesenen Stetigkeit von $\overline{I}_c|\overline{\mathfrak{R}}_c$ ist also $\sup \mu_c(Q_n) = \sup \overline{I}_c(\chi_{Q_n}) = I_c(\chi_Q) = \mu_c(Q)$. Folglich ist $\mu_c|\mathfrak{Q}_c$ und damit auch $\mu_c|\mathfrak{Q}$ σ-additiv.

Für den Nachweis, daß $\mu_p|\mathfrak{Q}$ (und damit übrigens auch $\mu_p|\mathfrak{Q}_p$) rein-endlich-additiv ist, sei $\varrho|\mathfrak{Q}$ eine σ-additive Funktion mit $0 \leq \varrho(Q) \leq \mu_p(Q)$ für alle $Q \in \mathfrak{Q}$. Für eine beliebige Funktion f aus der Menge $\mathfrak{E}$ der $\mathfrak{Q}$-Treppenfunktionen und eine Darstellung von f in der Form $f = \sum_{\nu=1}^{n} \lambda_\nu \chi_{Q_\nu}$ mit $Q_\nu \in \mathfrak{Q}$ und reellen Zahlen λ_ν setzen wir $R(f) = \sum_{\nu=1}^{n} \lambda_\nu \varrho(Q_\nu)$. Dann ist bekanntlich[12]) $R|\mathfrak{E}$ eindeutig definiert und ein stetiges, positives, lineares Funktional. Es ist $0 \leq R(f) \leq \overline{I}_p(f)$ für alle $f \geq 0$ aus $\mathfrak{E}$. Aus dieser Ungleichung folgt $\mathfrak{S}_p \subseteq \mathfrak{B}(\mathfrak{E}, R)$; also kann R nach 1.2 auf genau eine Weise zu einem positiven, linearen Funktional $\overline{R}$ auf $\mathfrak{S}_p$ fortgesetzt werden. Nach dem zu I. Bewiesenen (angewandt auf $R|\mathfrak{E}$ an Stelle von $I_c|\mathfrak{R}$) ist $\overline{R}|\mathfrak{S}_p$ stetig; ferner gilt $0 \leq \overline{R}(f) \leq \overline{I}_p(f)$ für alle $f \geq 0$ aus $\mathfrak{S}_p$. Wie im Beweis des Satzes 3 zeigt man, daß $\overline{R}|\mathfrak{S}_p$ auf genau eine Weise zu einem $\mathfrak{S}_p$-singulären, positiven, linearen Funktional $I_\varrho|\overline{\mathfrak{R}}_p$ fortgesetzt werden kann. Für jede Funktion $f \geq 0$ aus $\overline{\mathfrak{R}}_p$ gilt dann wegen (12) $I_\varrho(f) = \sup(\overline{R}(g); g \leq f, g \in \mathfrak{S}_p)$; hieraus folgt $0 \leq I_\varrho(f) \leq \overline{I}_p(f)$ für jedes solche f. Da $\overline{I}_p|\overline{\mathfrak{R}}_p$ nach dem zu I. Bewiesenen rein-unstetig ist, folgt also $I_\varrho(f) = 0$ für alle $f \in \overline{\mathfrak{R}}_p$ und hieraus $\varrho(Q) = 0$ für alle $Q \in \mathfrak{Q}$, falls I_ϱ ein stetiges Funktional ist. Um zu zeigen, daß $\mu_p|\mathfrak{Q}$ rein-endlich-additiv ist, haben wir also noch die Stetigkeit von I_ϱ zu beweisen. Hierzu sei $\{f_n\}_{n=1,2,\ldots}$ eine Folge nicht negativer Funktionen aus $\overline{\mathfrak{R}}_p$ mit $f_n \uparrow f$ und $f \in \overline{\mathfrak{R}}_p$. Wegen der Positivität von I_ϱ ist nur $I_\varrho(f) \leq \sup I_\varrho(f_n)$ zu zeigen. Setzen wir $g_n = \min(f_n, g)$ für eine beliebige Funktion $g \in \mathfrak{S}_p$ mit $0 \leq g \leq f$, so sind, da $\mathfrak{S}_p$ ein gesättigter, linearer Unterraum von $\overline{\mathfrak{R}}_p$ ist, alle g_n Elemente von $\mathfrak{S}_p$; ferner gilt $g_n \uparrow g$. Wegen der Stetigkeit von $\overline{R}|\mathfrak{S}_p$ ist also $\overline{R}(g) = \sup \overline{R}(g_n)$. Es ist $\overline{R}(g_n) \leq I_\varrho(f_n)$ für jedes n und somit $\overline{R}(g) \leq \sup I_\varrho(f_n)$. Da $g \in \mathfrak{S}_p$ beliebig mit $0 \leq g \leq f$ gewählt war, folgt hieraus die behauptete Ungleichung $I_\varrho(f) \leq \sup I_\varrho(f_n)$.

---

[12]) Vgl. etwa AUMANN [2], S. 365, Satz 9.1.3.1.

Aus der für alle $f \in \mathfrak{R}$ gültigen Gleichung $I = I_c + I_p$ folgt schließlich $\mu(Q) = \mu_c(Q) + \mu_p(Q)$ für alle $Q \in \mathfrak{Q}$. Also ist auf $\mathfrak{Q}$ μ_c der σ-additive und μ_p der rein-endlich-additive Teil von μ.

Korollar. *Auf $\mathfrak{S}$ ist $f \to \int f\,d\mu_c$ der stetige und $f \to \int f\,d\mu_p$ der rein-unstetige Teil des auf $\mathfrak{S}$ definierten Funktionals $f \to \int f\,d\mu$.*

Beweis. Es ist $\int f\,d\mu = \int f\,d\mu_c + \int f\,d\mu_p$ für jede Funktion $f \in \mathfrak{S}$, da auf $\mathfrak{R}$ die Gleichung $I = I_c + I_p$ gilt. Das auf $\mathfrak{S}$ definierte, positive, lineare Funktional $f \to \int f\,d\mu_c$ ist als Verengerung des stetigen Funktionals $I_c | \mathfrak{R}$ selbst stetig. Um zu zeigen, daß das auf $\mathfrak{S}$ definierte, positive, lineare Funktional $f \to \int f\,d\mu_p$ rein-unstetig ist, sei $K | \mathfrak{S}$ ein stetiges, positives, lineares Funktional mit $0 \leq K(f) \leq \int f\,d\mu_p$ für alle $f \geq 0$ aus $\mathfrak{S}$. Setzen wir $\varkappa(Q) = K(\chi_Q)$ für jedes $Q \in \mathfrak{Q}$, so ist $\varkappa | \mathfrak{Q}$ σ-additiv und es gilt $0 \leq \varkappa(Q) \leq \mu_p(Q)$ für alle $Q \in \mathfrak{Q}$. Nach der Behauptung II von Satz 4 ist also $\varkappa = 0$. Dann ist aber $K = 0$ auf $\mathfrak{S}$, da aus der für alle $f \geq 0$ aus $\mathfrak{S}$ gültigen Ungleichung $0 \leq K(f) \leq \int f\,d\mu_p$ folgt $\mathfrak{S} \subseteqq \mathfrak{B}(\mathfrak{E}, K)$ und somit die Fortsetzung von $K | \mathfrak{E}$ zu einem positiven, linearen Funktional auf $\mathfrak{B}(\mathfrak{E}, K)$ eindeutig bestimmt ist. Also ist das Funktional $f \to \int f\,d\mu_p$ rein-unstetig.

Nach Satz 3 sind die Funktionen aus $\mathfrak{R}$ in bezug auf jedes auf $\mathfrak{R}$ definierte, positive, lineare Funktional uneigentlich Riemann-integrierbar. Insbesondere existiert also $\overset{(u)}{\int} f\,d\mu_c$ für jedes $f \in \mathfrak{R}$. Über dieses auf $\mathfrak{R}$ definierte Funktional behaupten wir:

Satz 5. I. *Für jede Funktion $f \in \mathfrak{R}$ gilt*

$$I_c(f) = \overset{(u)}{\int} f\,d\mu_c. \tag{19}$$

II. *Das auf $\mathfrak{R}$ definierte, positive, lineare Funktional $f \to \int f\,d\mu$ hat $f \to \overset{(u)}{\int} f\,d\mu_c$ zum stetigen und $f \to \overset{(u)}{\int} f\,d\mu_p$ zum rein-unstetigen Teil.*

Beweis. *Zu* I. Es genügt die Gleichung $I_c(f) = \overset{(u)}{\int} f\,d\mu_c$ für jede nicht negative Funktion $f \in \mathfrak{R}$ zu beweisen. Es sei also $f \in \mathfrak{R}$ nicht negativ und außerdem vorerst beschränkt: $0 \leq f(x) \leq K$ für alle $x \in E$. Dann existiert eine Folge $\{g_n\}_{n=1,2,\ldots}$ von Funktionen aus $\mathfrak{S}$ mit $g_n \uparrow f$. Nach Satz 1 gibt es nämlich eine Folge $\{\alpha_n\}_{n=1,2,\ldots}$ reeller Zahlen mit $\alpha_1 > \alpha_2 > \cdots > 0$ und $\lim \alpha_n = 0$ derart, daß $Q_n = [f > \alpha_n]$ für jedes n I-quadrierbar ist. Die Folge der Funktionen $g_n = \min(f, K\chi_{Q_n})$ leistet dann das Verlangte (wegen (11) ist $g_n \in \mathfrak{S}$). Nach Satz 4, Beh. I ist I_c auf $\mathfrak{R}$ stetig; folglich gilt $I_c(f) = \sup I_c(g_n) = \sup \int g_n\,d\mu_c \leq \overset{(u)}{\int} f\,d\mu_c$. Da nach (13) andererseits $\overset{(u)}{\int} f\,d\mu_c \leq I_c(f)$ ist, ist damit die behauptete Gleichung bewiesen.

Nunmehr sei f eine beliebige nicht negative Funktion aus $\mathfrak{R}$. Dann ist $f_n = \min(n, f)$, $n = 1, 2, \ldots$, eine Folge beschränkter, nicht negativer Funktionen aus $\mathfrak{R}$ mit $f_n \uparrow f$; also gilt $I_c(f) = \sup I_c(f_n) = \sup \overset{(u)}{\int} f_n\,d\mu_c \leq \overset{(u)}{\int} f\,d\mu_c$ nach dem bisher Bewiesenen. Wegen (13) gilt also auch in diesem Falle (19).

Aus dem Beweis von Gl. (19) entnehmen wir noch das für das Folgende nützliche Ergebnis: Für jede Funktion $f \geq 0$ aus $\mathfrak{R}$ ist

$$\overset{(u)}{\int} f\, d\mu_c = \sup\left(\int g\, d\mu_c;\ g \leq f,\ g \in \mathfrak{S}\right). \tag{20a}$$

Zu II. Nach I. ist das auf $\mathfrak{R}$ definierte, positive, lineare Funktional $f \to \overset{(u)}{\int} f\, d\mu_c$ stetig. Das auf $\mathfrak{R}$ definierte, positive, lineare Funktional $f \to \overset{(u)}{\int} f\, d\mu_p$ ist rein-unstetig, da $0 \leq \overset{(u)}{\int} f\, d\mu_p \leq I_p(f)$ für jede Funktion $f \geq 0$ aus $\mathfrak{R}$ gilt. Es verbleibt also der Beweis der Gleichung $\overset{(u)}{\int} f\, d\mu = \overset{(u)}{\int} f\, d\mu_c + \overset{(u)}{\int} f\, d\mu_p$ für jedes $f \geq 0$ aus $\mathfrak{R}$. Hierzu zeigen wir, daß neben (20a) auch

$$\overset{(u)}{\int} f\, d\mu_p = \sup\left(\int g\, d\mu_p;\ g \leq f, g \in \mathfrak{S}\right) \tag{20b}$$

für alle $f \geq 0$ aus $\mathfrak{R}$ gilt. Die behauptete Gleichung folgt dann nämlich so: Es bezeichne Φ die steigend gerichtete Menge aller Funktionen $g \in \mathfrak{S}$ mit $g \leq f$; dann ist

$$\overset{(u)}{\int} f\, d\mu = \lim_{\Phi} \int g\, d\mu = \lim_{\Phi} \int g\, d\mu_c + \lim_{\Phi} \int g\, d\mu_p = \overset{(u)}{\int} f\, d\mu_c + \overset{(u)}{\int} f\, d\mu_p.$$

Wir haben also nur noch den Beweis von (20b) nachzutragen. Es sei hierzu $Q \in \mathfrak{Q}_c$ mit $\chi_Q \leq f$. Nach Satz 1 ist dann die Menge $Q' = [f > \alpha]$ für ein geeignetes α mit $0 < \alpha < 1$ Element von $\mathfrak{Q}$. Nach (11) ist $g = \min(f, \chi_{Q'})$ eine Funktion aus $\mathfrak{S}$; für diese gilt $\chi_Q \leq g \leq f$. – Aus der Definition der Riemann-integrierbaren Funktionen folgt unmittelbar die Gleichung

$$\overset{(u)}{\int} f\, d\mu_p = \sup\left(\int g\, d\mu_p;\ g \leq f, g \in \mathfrak{E}_p\right),$$

wenn $\mathfrak{E}_p$ die Menge der $\mathfrak{Q}_p$-Treppenfunktionen bezeichnet. Die Gl. (20b) ist daher bewiesen, wenn gezeigt ist, daß zu beliebigem $g \in \mathfrak{E}_p$ mit $0 \leq g \leq f$ stets ein $g' \in \mathfrak{S}$ mit $g \leq g' \leq f$ existiert. Es kann g dargestellt werden in der Form $g = \sum_{\nu=1}^{n} \lambda_\nu \chi_{Q_\nu}$ mit paarweise fremden Mengen $Q_\nu \in \mathfrak{Q}_p$ und reellen Zahlen $\lambda_\nu > 0$. Wegen $\chi_{Q_\nu} \leq \frac{1}{\lambda_\nu} f$ existiert, wie bereits bewiesen wurde, eine Funktion $g_\nu \in \mathfrak{S}$ mit $\chi_{Q_\nu} \leq g_\nu \leq \frac{1}{\lambda_\nu} f$. Setzen wir also $g' = \max(\lambda_1 g_1, \ldots, \lambda_n g_n)$, so ist g' eine Funktion aus $\mathfrak{S}$, die das Verlangte leistet.

Damit ist der Satz 5 vollständig bewiesen.

Im folgenden Satz vergleichen wir die auf $\mathfrak{R}$ definierten Funktionale I und $f \to \overset{(u)}{\int} f\, d\mu$ hinsichtlich der sich aus ihnen ableitenden Riemann-Integrale.

Satz 6. I. *Die auf $\mathfrak{R}$ definierten, positiven, linearen Funktionale I und $f \to \overset{(u)}{\int} f\, d\mu$ führen zur gleichen Menge $\mathfrak{S}$ Riemann-integrierbarer Funktionen und zum gleichen Riemann-Integral über $\mathfrak{S}$.*

II. *Das auf $\mathfrak{R}$ definierte, positive, lineare Funktional $f \to D(f) = I(f) - \overset{(u)}{\int} f\, d\mu$ ist rein-unstetig. Bezüglich D sind genau diejenigen Funktionen auf E Riemann-*

integrierbar, welche beschränkt und im Komplement einer I-quadrierbaren Menge gleich Null sind. Für jede solche Funktion ist das Riemann-Integral bezüglich D gleich Null.

Beweis. *Zu* I. Wir setzen $L(f)=\overset{(u)}{\int} f\,d\mu$ für jedes $f\in\mathfrak{R}$. Dann sei $\mathfrak{Q}_L$ das System der L-quadrierbaren Mengen und $\mu_L|\mathfrak{Q}_L$ der sich aus L ableitende Inhalt. Es genügt zu zeigen, daß $\mathfrak{Q}=\mathfrak{Q}_L$ und $\mu=\mu_L$ ist. Zunächst ist $\mathfrak{Q}\subseteqq\mathfrak{Q}_L$, da nach (13) $0\leqq L(f)\leqq I(f)$ für alle $f\geqq 0$ aus $\mathfrak{R}$ gilt. Ferner ist $\mu(Q)=\mu_L(Q)$ für jede Menge $Q\in\mathfrak{Q}$, da einerseits $f\to\bar{I}(f)$ bzw. $f\to\overset{(u)}{\int} f\,d\mu$ die einzige Fortsetzung von I bzw. L zu einem positiven, linearen Funktional auf $\bar{\mathfrak{R}}$ und andererseits $\overset{(u)}{\int} f\,d\mu=\int f\,d\mu$ für alle $f\in\mathfrak{S}$, also insbesondere für $f=\chi_Q$ ist. Es sei Q nun eine Menge aus $\mathfrak{Q}_L$; dann gibt es zu beliebigem $\varepsilon>0$ Funktionen $g,h\in\mathfrak{R}$ mit $0\leqq g\leqq\chi_Q\leqq h$ und $L(h-g)\leqq\varepsilon$. Zu g existiert nach (9) ein $g'\in\mathfrak{S}$ mit $g'\leqq g$ und $L(g)-\int g'\,d\mu\leqq\varepsilon$. Ferner beweist man wie im Beweis der Behauptung II von Satz 5 die Existenz einer Funktion $h'\in\mathfrak{S}$ mit $\chi_Q\leqq h'\leqq h$. Also gilt $g'\leqq\chi_Q\leqq h'$ und $\int(h'-g')\,d\mu\leqq L(h)-\int g'\,d\mu\leqq L(h-g)+\varepsilon\leqq 2\varepsilon$. Hieraus folgt $Q\in\mathfrak{Q}$ wegen der Vollständigkeit des Funktionals $f\to\int f\,d\mu$ auf $\mathfrak{S}$. Also ist $\mathfrak{Q}=\mathfrak{Q}_L$ und $\mu=\mu_L$.

Zu II. Nach Satz 5 ist $D(f)=I_c(f)+I_p(f)-\overset{(u)}{\int} f\,d\mu_c-\overset{(u)}{\int} f\,d\mu_p=I_p(f)-\overset{(u)}{\int} f\,d\mu_p$; da also $0\leqq D(f)\leqq I_p(f)$ für alle $f\geqq 0$ aus $\mathfrak{R}$ gilt, ist D reinunstetig. Nach der letzten Ungleichung ist ferner die Vervollständigung $\bar{I}|\bar{\mathfrak{R}}$ von I Verengerung der Vervollständigung $\bar{D}|\mathfrak{B}(\mathfrak{R},D)$ von D, insbesondere ist $\bar{\mathfrak{R}}\subseteqq\mathfrak{B}(\mathfrak{R},D)$. Da die Fortsetzung von D zu einem linearen Funktional auf $\mathfrak{B}(\mathfrak{R},D)$ nur auf genau eine Weise möglich ist, gilt daher $\bar{D}(f)=\bar{I}_p(f)-\overset{(u)}{\int} f\,d\mu_p$ für jedes $f\in\bar{\mathfrak{R}}$. Also ist $\bar{D}(f)=0$ für jedes $f\in\mathfrak{S}$. Hieraus folgt, daß jede I-quadrierbare Menge Q und mit ihr jede Teilmenge Q' von Q D-quadrierbar ist; bezeichnet δ den zu D gehörigen Inhalt, so ist $\delta(Q')=\bar{D}(\chi_{Q'})\leqq\bar{D}(\chi_Q)=0$, also $\delta(Q')=0$ für jede solche Menge Q'. Ist umgekehrt Q' eine beliebige D-quadrierbare Menge, so existiert eine I-quadrierbare Menge Q mit $Q'\subseteqq Q$ und es gilt folglich wieder $\delta(Q')=0$. Zu Q' existiert nämlich ein $h\in\mathfrak{R}$ mit $\chi_{Q'}\leqq h$ und hierzu eine reelle Zahl α mit $0<\alpha<1$ derart, daß $Q=[h>\alpha]$ I-quadrierbar ist; diese Menge leistet das Verlangte. Somit ist das System der D-quadrierbaren Mengen gleich dem System der Teilmengen Q' beliebiger I-quadrierbarer Mengen Q; für jede D-quadrierbare Menge Q' gilt $\delta(Q')=0$. Hieraus folgt die Behauptung (z. B. nach Satz 2).

§ 3. Riemann-Integrierbarkeit bezüglich eines linearen Funktionals. Konkreter Fall

3.1. Die in den beiden vorhergehenden Paragraphen gewonnenen Ergebnisse wenden wir nun auf folgenden Spezialfall an:

Die Menge E sei ein *lokal kompakter*[13]), topologischer *Raum*. Die Elemente von $\mathfrak{R}$ seien auf E definierte, *stetige*, (endliche) *reelle Funktionen, welche*

[13]) Im folgenden bedeutet „kompakt" stets genauer „bikompakt und Hausdorffsch".

im Unendlichen verschwinden, d.h. für jede solche Funktion g und jede reelle Zahl $\alpha > 0$ ist die Urbildmenge $[|g| \geq \alpha]$ kompakt. Hieraus folgt insbesondere die Beschränktheit einer jeden Funktion $g \in \mathfrak{R}$. Die Funktionenmenge $\mathfrak{R}$ besitze wiederum die Eigenschaften (1) bis (3) und genüge daneben noch folgenden zwei Zusatzforderungen:

(21) *Zu jedem Punkt* $x \in E$ *existiert eine Funktion* $g \in \mathfrak{R}$ *mit* $g(x) \neq 0$.

(22) *Zu je zwei verschiedenen Punkten* $x, y \in E$ *existiert eine Funktion* $h \in \mathfrak{R}$ *mit* $h(x) \neq h(y)$.

Beispielsweise genügt die Menge $\mathfrak{R} = \mathfrak{R}(E)$ aller stetigen, reellen Funktionen $g|E$ mit kompaktem Träger[14]) den Bedingungen (1) bis (3) sowie (21) und (22).

Aus den genannten Eigenschaften von $\mathfrak{R}$ folgt

Lemma 4. *Zu jeder stetigen, im Unendlichen verschwindenden, reellen Funktion* $f|E$ *existiert eine gleichmäßig gegen* f *konvergente Folge* $\{g_n\}_{n=1,2,\ldots}$ *von Funktionen aus* $\mathfrak{R}$ *mit der Eigenschaft* (4).

Beweis. Die Menge $\mathfrak{A}$ der auf E definierten, durch Funktionen aus $\mathfrak{R}$ gleichmäßig approximierbaren Funktionen besteht aus stetigen, im Unendlichen verschwindenden Funktionen und ist nach Lemma 1 eine Algebra. Wegen $\mathfrak{R} \subseteq \mathfrak{A}$ besitzt $\mathfrak{A}$ die Eigenschaften (21) und (22) von $\mathfrak{R}$. Nach einem Approximationssatz von NAKANO [*15*][15]) kann daher jede stetige, im Unendlichen verschwindende Funktion auf E durch Funktionen aus $\mathfrak{A}$ gleichmäßig approximiert werden. Da $\mathfrak{A}$ bezüglich gleichmäßiger Konvergenz abgeschlossen ist, so ist also $\mathfrak{A}$ das System aller stetigen, im Unendlichen verschwindenden Funktionen auf E. Die Behauptung folgt daher aus Lemma 1.

3.2. Neben E und $\mathfrak{R}$ sei wie bisher ein auf $\mathfrak{R}$ definiertes, positives, lineares Funktional I vorgegeben. Gemäß § 1 leiten sich aus I ab: Das System $\mathfrak{Q}$ der I-quadrierbaren Teilmengen von E, der Inhalt $\mu|\mathfrak{Q}$ und das System $\mathfrak{S}$ aller bezüglich I Riemann-integrierbaren Funktionen.

Wir behaupten von $\mu|\mathfrak{Q}$ die folgenden vier Eigenschaften: (1) jede Menge $Q \in \mathfrak{Q}$ ist relativ kompakt, d.h. es ist $\overline{Q}$[16]) kompakt; (2) in $\mathfrak{Q}$ ist eine offene Basis von E enthalten; (3) mit Q sind auch $\overline{Q}$ und $\underline{Q}$ Elemente von $\mathfrak{Q}$ und es ist $\mu(Q) = \mu(\overline{Q}) = \mu(\underline{Q})$; (4) zu jeder Menge $Q \in \mathfrak{Q}$ und jeder reellen Zahl $\varepsilon > 0$ existiert eine offene Menge $U \in \mathfrak{Q}$ mit $Q \subseteq U$ und $\mu(U - Q) \leq \varepsilon$.

Nach HAUPT und PAUC [*11*] sind bei Gültigkeit von (1) und (2) die beiden Eigenschaften (3) und (4) äquivalent mit den beiden folgenden: (3') für jedes $Q \in \mathfrak{Q}$ ist $\overline{Q} - \underline{Q} \in \mathfrak{Q}$; (4') zu jeder Menge $Q \in \mathfrak{Q}$ und zu jedem $\varepsilon > 0$ existiert eine offene Menge $U \in \mathfrak{Q}$ mit $\overline{Q} - \underline{Q} \subseteq U$ und $\mu(U) \leq \varepsilon$. Ein Inhalt $\mu|\mathfrak{Q}$ mit den Eigenschaften (1), (2), (3') und (4') heißt nach HAUPT-PAUC [*11*] an

[14]) Als *Träger* einer reellen Funktion $g|E$ bezeichnet man die abgeschlossene Hülle der Menge $[g \neq 0]$.

[15]) Vgl. hierzu auch NÖBELING und BAUER [*17*], S. 72.

[16]) Mit $\overline{A}$ bezeichnen wir die abgeschlossene Hülle, mit $\underline{A}$ den offenen Kern einer Menge $A \subseteq E$.

(die Topologie von) E *adaptiert*. Unsere Behauptung kann also auch so ausgesprochen werden:

Satz 7. *Der Inhalt* $\mu|\mathfrak{Q}$ *ist an* E *adaptiert.*

Beweis. Wir zeigen, daß $\mu|\mathfrak{Q}$ die Eigenschaften (1) bis (4) besitzt. — *Zu* (1). Zu einer Menge $Q\in\mathfrak{Q}$ existiert eine Funktion $h\in\mathfrak{R}$ mit $\chi_Q\leq h$; folglich ist $Q\subseteqq[h\geq 1]$. Da h stetig ist und im Unendlichen verschwindet, ist $[h\geq 1]$ kompakt, also Q relativ kompakt. — *Zu* (2). Es sei V eine offene Menge in E und x_0 ein Punkt von V. Zu zeigen ist die Existenz einer offenen Menge $U\in\mathfrak{Q}$ mit $x_0\in U$ und $U\subseteqq V$. Zu x_0 und V existiert eine stetige Funktion $f|E$ mit $0\leq f\leq 1$, $f(x_0)=1$ und in V enthaltenem kompakten Träger (vgl. [8], S. 49). Nach Lemma 4 gibt es zu f eine Funktion $g\in\mathfrak{R}$ mit $0\leq g\leq f\leq g+\frac{1}{2}$ und weiter nach Satz 1 eine reelle Zahl α mit $0<\alpha<\frac{1}{2}$, für welche die offene Urbildmenge $U=[g>\alpha]$ ein Element von $\mathfrak{Q}$ ist. Aus $f(x_0)=1$ folgt $g(x_0)\geq\frac{1}{2}$; also ist $x_0\in U$. Für einen beliebigen Punkt $x\in U$ ist $f(x)\geq g(x)>\alpha>0$; da f im Komplement von V gleich Null ist, gilt also $x\in V$. Folglich ist $U\subseteqq V$. Die Menge U leistet somit das Verlangte. — *Zu* (3). Es sei Q eine Menge aus $\mathfrak{Q}$, g und h seien Funktionen aus $\mathfrak{R}$ mit $g\leq\chi_Q\leq h$. Wegen der Stetigkeit von g und h gilt dann auch $g\leq\chi_{\mathring{Q}}\leq\chi_Q\leq\chi_{\overline{Q}}\leq h$. Hieraus und aus der Definition von μ folgt, daß $\overline{Q}$ und $\mathring{Q}$ quadrierbare Mengen mit $\mu(\overline{Q})=\mu(\mathring{Q})=\mu(Q)$ sind. — *Zu* (4). Es sei Q eine quadrierbare Menge. Zu einer beliebig reellen Zahl η mit $0<\eta<1$ existiert dann ein $h\in\mathfrak{R}$ mit $\chi_Q\leq h$ und $I(h)\leq\mu(Q)+\eta$. Nach Satz 1 gibt es eine Zahl η' mit $0<\eta'\leq\eta$ derart, daß die offene Urbildmenge $U=[h>1-\eta']$ quadrierbar ist. Wir behaupten, daß $Q\subseteqq U$ und $\mu(U-Q)\leq\frac{\eta}{1-\eta}\big(1+\mu(Q)\big)$ ist; hieraus folgt dann (4). Aus $x\in Q$ folgt $h(x)\geq 1>1-\eta'$ und hieraus $x\in U$; also ist $Q\subseteqq U$. Nach Definition von U ist $(1-\eta')\,\chi_U\leq h$ und daher $(1-\eta')\mu(U)\leq I(h)$. Hieraus ergibt sich $(1-\eta)\mu(U)\leq(1-\eta')\mu(U)\leq I(h)\leq\mu(Q)+\eta$ und weiter $(1-\eta)\mu(U-Q)=(1-\eta)\big(\mu(U)-\mu(Q)\big)\leq\mu(Q)+\eta-(1-\eta)\,\mu(Q)=\eta\big(1+\mu(Q)\big)$. Es gilt also $\mu(U-Q)\leq\frac{\eta}{1-\eta}\big(1+\mu(Q)\big)$, wie behauptet. — Damit ist der Satz bewiesen.

Allein aus der Adaption von $\mu|\mathfrak{Q}$ an E ergeben sich nun unter anderem die im folgenden Korollar genannten Eigenschaften von μ und $\mathfrak{Q}$.

Korollar. I. *Zu jeder kompakten Menge* K *und jeder offenen Menge* U *mit* $K\subseteqq U$ *existiert eine Menge* $Q\in\mathfrak{Q}$ *mit* $K\subseteqq Q\subseteqq U$.

II. *Zu jeder Menge* $Q\in\mathfrak{Q}$ *sowie zu beliebigem* $\varepsilon>0$ *existieren eine kompakte Menge* $K\in\mathfrak{Q}$ *und eine offene Menge* $U\in\mathfrak{Q}$ *mit* $K\subseteqq Q\subseteqq U$ *und* $\mu(U-K)\leq\varepsilon$.

Beweis. Vgl. HAUPT-PAUC [11], S. 192.

3.3. Nach 3.2 genügt $\mathfrak{Q}$ den bei BOURBAKI [8], S. 164—165 genannten Bedingungen (PC_I), (PC_{II}) und $\mu|\mathfrak{Q}$ den Bedingungen (PM_I), (PM_{II}). Folglich existiert nach Bourbaki *genau ein* positives Radonsches Maß ϱ auf E, in bezug auf welches alle Mengen $Q\in\mathfrak{Q}$ ϱ-integrierbar mit $\varrho(Q)=\mu(Q)$ sind.

Dieses eindeutig bestimmte Radonsche Maß läßt sich auch wie folgt kennzeichnen[17]):

Satz 8. *Jede auf E stetige Funktion mit kompaktem Träger ist Riemann-integrierbar bezüglich I. Das somit auf $\mathfrak{K}(E)$ definierte lineare Funktional $f \to \int f\,d\mu$ ist das einzige positive Radonsche Maß ϱ auf E, in bezug auf welches jede Menge $Q \in \mathfrak{Q}$ ϱ-integrierbar mit $\varrho(Q) = \mu(Q)$ ist. Allgemeiner ist jede bezüglich I Riemann-integrierbare Funktion f sowohl ϱ-integrierbar als auch bezüglich ϱ Riemann-integrierbar*[18]) *und es ist*

$$\int f\,d\varrho = \int f\,d\mu; \tag{23}$$

in dieser Gleichung kann $\int f\,d\varrho$ sowohl als ϱ-Integral als auch als Riemann-Integral bezüglich ϱ gedeutet werden.

Beweis. Es sei f eine Funktion aus $\mathfrak{K}(E)$ und K der kompakte Träger von f. Nach dem Korollar zu Satz 7 gibt es dann eine Menge $Q \in \mathfrak{Q}$ mit $K \subseteq Q$. Nach Lemma 4 kann f durch Funktionen aus $\mathfrak{K}$ gleichmäßig approximiert werden. Also folgt aus dem Korollar 2 zu Satz 2 und aus Lemma 2, daß f Riemann-integrierbar bezüglich I ist. Setzen wir $\int f\,d\varrho = \int f\,d\mu$ für jedes $f \in \mathfrak{K}(E)$, so ist offenbar ϱ ein positives Radonsches Maß auf E. Daß jede Menge $Q \in \mathfrak{Q}$ ϱ-integrierbar mit $\varrho(Q) = \mu(Q)$ ist, ergibt sich aus der jetzt zu beweisenden, allgemeineren Aussage: Jede Funktion $f \in \mathfrak{S}$ ist ϱ-integrierbar und zugleich bezüglich ϱ Riemann-integrierbar; es gilt (23). Hierzu kann ohne Beschränkung der Allgemeinheit $f \geqq 0$ vorausgesetzt werden. Da f im Komplement einer I-quadrierbaren Menge gleich Null ist, existiert nach

[17]) Bezüglich der im folgenden verwendeten Begriffe aus der Theorie Radonscher Maße sei auf BOURBAKI [8] verwiesen. Ein positives Radonsches Maß ϱ auf E ist demnach ein positives, lineares Funktional auf der Menge $\mathfrak{K}(E)$ aller stetigen Funktionen $f|E$ mit kompaktem Träger, wobei $\int f\,d\varrho$ für $\varrho(f)$ steht und den Funktionswert von ϱ an der Stelle $f \in \mathfrak{K}(E)$ bezeichnet. Nach BOURBAKI definiert man für jede nicht negative Funktion $q|E$ das ϱ-Oberintegral $\int^* q\,d\varrho$. Es erweist sich dann $N^*(f) = \int^* |f|\,d\varrho$ als Quasi-Norm im System aller Funktionen $f|E$. Eine Funktion $f|E$ heißt ϱ-integrierbar, wenn $N^*(f) < +\infty$ ist und eine Folge von Funktionen $f_n \in \mathfrak{K}(E)$ mit $\lim_{n\to\infty} N^*(f - f_n) = 0$ existiert; man setzt $\int f\,d\varrho = N^*(f^+) - N^*(f^-)$. Eine Teilmenge A von E, deren charakteristische Funktion χ_A ϱ-integrierbar ist, wird selbst als ϱ-integrierbar bezeichnet; man setzt $\varrho(A) = \int \chi_A\,d\varrho$. Nachträglich erweist sich dann $\int f\,d\varrho$ als Unterteilungsintegral nach dieser Mengenfunktion. — Wendet man dagegen auf das auf $\mathfrak{K}(E)$ definierte Funktional $f \to \int f\,d\varrho$ den sog. ersten (mit Folgen und nicht mit gerichteten Systemen arbeitenden) Erweiterungsprozeß von STONE [21] (vgl. auch [12b], S. 130ff.) an, so gewinnt man ein für alle nicht negativen Funktionen $q|E$ erklärtes „*schwaches*" ϱ-Oberintegral $\overline{\int} q\,d\varrho$. Aus diesem leiten sich analog eine Quasi-Norm $N(f) = \overline{\int} |f|\,d\varrho$ über dem System aller Funktionen $f|E$ und ein Integralbegriff ab. Jede im Sinne von STONE ϱ-integrierbare Funktion ist auch im Sinne von BOURBAKI ϱ-integrierbar und die zugehörigen Integrale sind einander gleich. Wenn im folgenden von einer Funktion oder Menge (d.h. also von deren charakteristischen Funktionen) die ϱ-Integrierbarkeit behauptet wird, so kann hierbei, ohne daß dies eigens betont wird, „ϱ-integrierbar" sowohl im Sinne von STONE als auch im Sinne von BOURBAKI verstanden werden. Die Beweise sind dementsprechend angelegt.

[18]) „Bezüglich ϱ Riemann-integrierbar" soll genauer heißen „bezüglich des auf $\mathfrak{K}(E)$ definierten Funktionals $f \to \int f\,d\varrho$ Riemann-integrierbar".

Lemma 2 eine nicht negative Funktion $h_0 \in \mathfrak{R}$ derart, daß für jedes $x \in E$ aus $h_0(x) < 1$ folgt $f(x) = 0$. Nach Lemma 4, angewandt auf $\mathfrak{R}(E)$ an Stelle von $\mathfrak{R}$, folgt die Existenz einer Funktion $h_0' \in \mathfrak{R}(E)$ mit $0 \leq h_0' \leq 2\, h_0 \leq h_0' + 1$. Auch diese Funktion besitzt dann die Eigenschaft, daß aus $h_0'(x) < 1$ stets $f(x) = 0$ folgt. Da f in der Vervollständigung $\overline{\mathfrak{R}}$ von $\mathfrak{R}$ bezüglich I liegt, gibt es zu jeder reellen Zahl $\varepsilon > 0$ Funktionen $g, h \in \mathfrak{R}$ mit $0 \leq g \leq f \leq h$ und $I(h-g) \leq \varepsilon$. Mit f besitzt auch g einen kompakten Träger, d.h. es ist $g \in \mathfrak{R}(E)$. Durch erneute Anwendung von Lemma 4 ergibt sich die Existenz einer Funktion $h' \in \mathfrak{R}(E)$ mit $0 \leq h' \leq h \leq h' + \varepsilon$. Setzt man $g'' = g$ und $h'' = h' + \varepsilon h_0'$, so sind g'' und h'' Funktionen aus $\mathfrak{R}(E)$, nach Wahl von h_0' gilt $g'' \leq f \leq h''$ und es ist $\int (h''-g'')\, d\varrho = \int (h'-g)\, d\mu + \varepsilon \int h_0'\, d\mu \leq I(h-g) + 2\varepsilon I(h_0) \leq \varepsilon\big(1 + 2I(h_0)\big)$. Also ist f bezüglich ϱ Riemann-integrierbar. Wegen $\int g''\, d\varrho = \int g''\, d\mu$ und $\int h''\, d\varrho = \int h''\, d\mu$ gilt (23) in dem Sinne, daß $\int f\, d\varrho$ das Riemann-Integral von f bezüglich ϱ bedeutet. Nun ist aber $\int (f-g'')\, d\varrho \leq \int (h''-g'')\, d\varrho$ und $g'' \in \mathfrak{R}(E)$; also ist f auch ϱ-integrierbar und in (23) kann $\int f\, d\varrho$ auch als ϱ-Integral gedeutet werden.

Korollar. *Die Funktionale $I|\mathfrak{R}$ und $\int f\, d\varrho\,|\,\mathfrak{R}(E)$ führen zum gleichen an E adaptierten Inhalt, zur gleichen Menge $\mathfrak{S}$ Riemann-integrierbarer Funktionen und zum gleichen Riemann-Integral über $\mathfrak{S}$. Es ist $\mathfrak{S}$ gleich der Vervollständigung von $\mathfrak{R}(E)$ bezüglich ϱ.*

Beweis. Bezeichnet wie bisher $\mathfrak{S}$ die Menge der bezüglich I Riemann-integrierbaren Funktionen, so ist nach Satz 8 $\mathfrak{R}(E) \subseteq \mathfrak{S}$ und $\int f\, d\varrho = \int f\, d\mu$ für jede Funktion $f \in \mathfrak{R}(E)$. Da $\mathfrak{S}$ bezüglich des über $\mathfrak{S}$ definierten Funktionals $f \to \int f\, d\mu$ vollständig ist, ergibt sich hieraus $\overline{\mathfrak{R}(E)} \subseteq \mathfrak{S}$, wenn $\overline{\mathfrak{R}(E)}$ die Vervollständigung von $\mathfrak{R}(E)$ bezüglich ϱ bezeichnet. Es gilt sogar $\overline{\mathfrak{R}(E)} = \mathfrak{S}$, da nach Satz 8 jede Funktion aus $\mathfrak{S}$ auch bezüglich ϱ Riemann-integrierbar, also jedenfalls ein Element von $\overline{\mathfrak{R}(E)}$ ist. Aus der Gleichung $\overline{\mathfrak{R}(E)} = \mathfrak{S}$ und Satz 8 folgt aber der Rest der Behauptung.

Bemerkung. Nach dem zuletzt bewiesenen Korollar bedeutet es keine Beschränkung der Allgemeinheit, die Theorie des Riemann-Integrals in lokal kompakten Räumen nur für Radonsche Maße zu entwickeln, wie dies Haupt und Pauc in [10], [19], [12b], S. 174ff., für kompakte und in einer in Vorbereitung befindlichen Arbeit auch für lokal kompakte Räume getan haben. Ausgangspunkt der Untersuchungen von Haupt-Pauc ist dabei allerdings ein an den Raum adaptierter Inhalt. Vermöge der Bemerkung am Beginn der Nr. 3.3 entsprechen sich aber adaptierte Inhalte und positive Radonsche Maße in umkehrbar eindeutiger Weise.

3.4. Da Riemann-Integrierbarkeit bezüglich I dasselbe wie Riemann-Integrierbarkeit bezüglich ϱ bedeutet, ist zu erwarten, daß sich die von Pauc [19] für den Fall eines kompakten Raumes angegebene Kennzeichnung der Riemann-integrierbaren Funktionen als fast überall stetige Funktionen auch auf den Fall eines lokal kompakten Raumes übertragen läßt.

Satz 9. *I-quadrierbar sind genau die relativ kompakten Teilmengen Q von E, deren Begrenzung $\overline{Q} - \underline{Q}$ eine ϱ-Nullmenge ist. Riemann-integrierbar*

bezüglich I sind genau die beschränkten, ϱ-fast überall stetigen Funktionen auf E mit kompaktem Träger[19]).

Beweis. Nach der Adaptionseigenschaft (1) von S. 462 ist eine I-quadrierbare Menge Q stets relativ kompakt; nach der Adaptionseigenschaft (3) ist $\overline{Q}-\underline{Q}\in\mathfrak{Q}$ und $\mu(\overline{Q}-\underline{Q})=0$. Aus Satz 8 folgt daher, daß $\overline{Q}-\underline{Q}$ eine ϱ-Nullmenge ist. Ist umgekehrt Q eine relativ kompakte Menge mit $\varrho(\overline{Q}-\underline{Q})=0$, so existiert zu jedem $\varepsilon>0$ eine offene Menge $U\in\mathfrak{Q}$ mit $\overline{Q}-\underline{Q}\subseteqq U$ und $\mu(U)=\varrho(U)\leq\varepsilon$. Wegen der Vollständigkeit des Inhalts μ folgt hieraus nach HAUPT-PAUC [*11*], S. 193 die I-Quadrierbarkeit von Q. Damit ist die erste Hälfte der Behauptung bewiesen. — Nunmehr sei f eine bezüglich I Riemann-integrierbare Funktion. Dann ist jedenfalls f beschränkt, ferner besitzt f einen kompakten Träger, da f im Komplement einer Menge $Q\in\mathfrak{Q}$ gleich Null und Q relativ kompakt ist. Es bezeichne $w(x)$ die Schwankung von f in einem Punkte $x\in E$; für ein $\gamma>0$ sei F_γ die Urbildmenge $[w\geq\gamma]$. Zu beliebigem $\varepsilon>0$ existieren paarweise fremde Mengen $Q_1,\ldots,Q_n\in\mathfrak{Q}$ sowie reelle Zahlen α_ν,β_ν $(\nu=1,\ldots,n)$ mit $\sum_{\nu=1}^{n}\alpha_\nu\chi_{Q_\nu}\leq f\leq\sum_{\nu=1}^{n}\beta_\nu\chi_{Q_\nu}$ und $\sum_{\gamma=1}^{n}(\beta_\nu-\alpha_\nu)\,\mu(Q_\nu)\leq\varepsilon$. Es sei J die Menge aller Indizes $\nu=1,\ldots,n$, für welche der offene Kern $\underline{Q}_\nu$ von Q_ν mit F_γ einen nicht leeren Durchschnitt hat. Für jeden Punkt $x\in\underline{Q}_\nu\cap F_\gamma$ ist dann $\underline{Q}_\nu$ eine Umgebung von x und daher $\beta_\nu-\alpha_\nu\geq w(x)\geq\gamma$; also gilt $\beta_\nu-\alpha_\nu\geq\gamma$ für jedes $\nu\in J$. Hieraus folgt $\sum_{\nu\in J}\mu(Q_\nu)\leq\frac{1}{\gamma}\sum_{\nu\in J}(\beta_\nu-\alpha_\nu)\,\mu(Q_\nu)\leq\frac{\varepsilon}{\gamma}$.

Ferner gilt $F_\gamma\subseteqq\bigcup_{\nu=1}^{n}\overline{Q}_\nu$, da zu jedem Punkt $x\notin\bigcup_{\nu=1}^{n}\overline{Q}_\nu$ eine Umgebung U von x mit $U\cap\bigcup_{\nu=1}^{n}\overline{Q}_\nu=\emptyset$ existiert und somit $w(x)=0$ ist. Wir erhalten daher die Relation $F_\gamma-\bigcup_{\nu\in J}Q_\nu\subseteqq N$, wobei $N=\bigcup_{\nu=1}^{n}(\overline{Q}_\nu-\underline{Q}_\nu)$ nach dem ersten Teil des Beweises I-quadrierbar mit $\mu(N)=\varrho(N)=0$ ist. Setzen wir $Q_0=\bigcup_{\nu\in J}Q_\nu$, so ist $F_\gamma\subseteqq Q_0\cup N$ und $\mu(Q_0\cup N)=\sum_{\nu\in J}\mu(Q_\nu)\leq\frac{\varepsilon}{\gamma}$. Wegen der Vollständigkeit des Inhalts $\mu|\mathfrak{Q}$ ist also $F_\gamma\in\mathfrak{Q}$ und $\mu(F_\gamma)=\varrho(F_\gamma)=0$. Da für die Menge F aller Unstetigkeitsstellen von f die Gleichung $F=\bigcup_{n=1}^{\infty}F_{1/n}$ gilt, ferner stets $F_{1/n}\subseteqq F_{1/n+1}$ ist, so ist F eine ϱ-Nullmenge, also f ϱ-fast überall stetig. — Umgekehrt sei $f|E$ beschränkt, ϱ-fast überall stetig und im Komplement einer kompakten Menge K gleich Null. Nach dem Korollar zu Satz 7 kann ohne Beschränkung der Allgemeinheit K als I-quadrierbar vorausgesetzt werden. Für ein $\varepsilon>0$ sei F_ε wie oben als Urbildmenge $[w\geq\varepsilon]$ definiert. Da w oberhalb stetig ist, so ist F_ε abgeschlossen und wegen $F_\varepsilon\subseteqq K$ sogar kompakt. Somit ist $\varrho(F_\varepsilon)=0$ und damit auch die Begrenzung von F_ε eine ϱ-Nullmenge,

[19]) In Übereinstimmung mit dem in [17]) Gesagten ist hierbei „ϱ-fast überall stetig" im Sinne der BOURBAKIschen Theorie zu verstehen und heißt: es gilt $\varrho^*(F)=\int^*\chi_F\,d\varrho=0$ für die Menge F aller Unstetigkeitsstellen der betreffenden Funktion f. Der folgende Beweis zeigt, daß F dann sowohl im Sinne von BOURBAKI als auch von STONE ϱ-integrierbar und $\varrho(F)=0$ ist.

da f ϱ-fast überall stetig ist; nach dem bereits bewiesenen ersten Teil der Behauptung ist dann F_ε sogar I-quadrierbar und nach Satz 8 $\mu(F_\varepsilon)=\varrho(F_\varepsilon)=0$. Folglich existiert eine offene Menge $U_0\in\mathfrak{Q}$ mit $\mu(U_0)\leq\varepsilon$ und $F_\varepsilon\subseteq U_0$. Für jeden Punkt $x\in K-F_\varepsilon$ ist $w(x)<\varepsilon$; daher existiert eine Umgebung $U_x\in\mathfrak{Q}$ von x, in welcher f um höchstens ε schwankt. Wegen der Kompaktheit von K gibt es endlich viele Punkte $x_1,\ldots,x_n\in K-F_\varepsilon$ derart, daß K von den I-quadrierbaren, offenen Mengen $U_0, U_1=U_{x_1},\ldots,U_n=U_{x_n}$ überdeckt wird. Setzen wir $Q_0=U_0\cap K$, $Q_1=U_1\cap K-Q_0$ und allgemein $Q_\nu=U_\nu\cap K-\bigcup_{\varrho=0}^{\nu-1}Q_\varrho$, so sind die Mengen $Q_0,\ldots,Q_n$ paarweise fremd und I-quadrierbar, ferner ist $K=\bigcup_{\nu=0}^{n}Q_\nu$. Bezeichnet α_ν das Infimum und β_ν das Supremum von f auf Q_ν, $\nu=0,1,\ldots,n$, so gilt $\sum_{\nu=0}^{n}\alpha_\nu\chi_{Q_\nu}\leq f\leq\sum_{\nu=0}^{n}\beta_\nu\chi_{Q_\nu}$ und $\sum_{\nu=0}^{n}(\beta_\nu-\alpha_\nu)\mu(Q_\nu)\leq 2M\varepsilon+\varepsilon\sum_{\nu=1}^{n}\mu(Q_\nu)\leq\varepsilon\big(2M+\mu(K)\big)$; hierbei bedeutet M eine obere Schranke für den Betrag von f auf E. Also ist f Riemann-integrierbar bezüglich I.

3.5. Bislang blieb eine Folgerung aus der Adaption des Inhalts $\mu|\mathfrak{Q}$ an E unberücksichtigt. Nach HAUPT-PAUC [11] gilt nämlich: *Ein an E adaptierter Inhalt ist stets σ-additiv.* Es ist also

$$\mu_c=\mu \quad\text{und}\quad \mu_p=0 \tag{24}$$

der σ-additive bzw. rein-endlich-additive Teil von μ. Vermöge der in 2.3 aufgezeigten Koppelung der Zerlegung $\mu=\mu_c+\mu_p$ mit der Zerlegung $I=I_c+I_p$ von I in seinen stetigen Teil I_c und rein-unstetigen Teil I_p lassen sich nunmehr über die Zerlegungsteile I_c und I_p genauere Aussagen machen. Wir benützen hierzu die in § 2 eingeführten Bezeichnungen.

Satz 10. I. *Jede Funktion $f\in\mathfrak{R}$ ist ϱ-integrierbar und uneigentlich Riemann-integrierbar; es gilt*

$$I_c(f)=\int f\,d\varrho=\overset{(u)}{\int} f\,d\mu. \tag{25}$$

II. *Es ist $\mathfrak{Q}_c=\mathfrak{Q}$ und $\mathfrak{S}_c=\mathfrak{S}$.*

III. *Es ist $\mathfrak{Q}_p$ das System aller relativ kompakten Teilmengen von E und $\mathfrak{S}_p$ das System aller beschränkten Funktionen $f|E$ mit kompaktem Träger.*

Beweis. *Zu* I.. Da $\mu_c=\mu$ ist, folgt aus Satz 5, Beh. I die Gültigkeit der Gleichung $I_c(f)=\overset{(u)}{\int} f\,d\mu$ für alle $f\in\mathfrak{R}$. Zu zeigen ist also noch, daß jede Funktion $f\in\mathfrak{R}$ ϱ-integrierbar und $\int f\,d\varrho=\overset{(u)}{\int} f\,d\mu$ ist; hierzu kann ohne Beschränkung der Allgemeinheit $f\geq 0$ vorausgesetzt werden. Nach dem Lemma 4 (angewandt auf $\mathfrak{R}(E)$ an Stelle von $\mathfrak{R}$) existiert eine nicht fallende Folge $\{g_n\}$ nicht negativer Funktionen aus $\mathfrak{R}(E)$ mit $\sup g_n=f$. Wegen $\int g_n\,d\varrho=\int g_n\,d\mu\leq\overset{(u)}{\int} f\,d\mu<+\infty$, $n=1,2,\ldots$, ist also f ϱ-integrierbar und $\int f\,d\varrho=\sup\int g_n\,d\varrho\leq\overset{(u)}{\int} f\,d\mu$. Für eine beliebige Funktion $g\in\mathfrak{S}$ mit $g\leq f$ ist nach

(23) $\int g\,d\mu = \int g\,d\varrho \leqq \int f\,d\varrho$ und folglich $\overset{(u)}{\int} f\,d\mu \leqq \int f\,d\varrho$. Also ist $\int f\,d\varrho = \overset{(u)}{\int} f\,d\mu$. — *Zu* II. Die Behauptung folgt aus I. und Satz 6, Beh. I. — *Zu* III. Da jede I-quadrierbare Menge relativ kompakt und jede kompakte Menge in einer I-quadrierbaren enthalten ist, folgt die Behauptung aus I. und Satz 6, Beh. II.

Der Satz 10 zeigt, daß es zu jedem *stetigen*, positiven, linearen Funktional $I|\mathfrak{R}$ ein eindeutig bestimmtes, positives Radonsches Maß ϱ auf E gibt, in bezug auf welches $I(f) = \int f\,d\varrho$ für alle Funktionen $f \in \mathfrak{R}$ gilt. Eine erste Kennzeichnung der *rein-unstetigen* Funktionale auf $\mathfrak{R}$ enthält das folgende

Korollar. *Gleichwertig sind folgende drei Aussagen:*

I. *Das Funktional $I|\mathfrak{R}$ ist rein-unstetig.*

II. *Es ist $I(f) = 0$ für jede Funktion $f \in \mathfrak{R}$ mit kompaktem Träger.*

III. *Der Inhalt $\mu|\mathfrak{Q}$ ist identisch Null.*

Beweis. *Aus* I. *folgt* II. Eine Funktion $f \in \mathfrak{R}$ mit kompaktem Träger ist nach Satz 8 Riemann-integrierbar; folglich gilt $\int f\,d\mu = \int f\,d\varrho = I(f)$. Nach Satz 10 gilt aber auch $\int f\,d\varrho = I_c(f)$. Wegen $I_c = 0$ ist also $I(f) = 0$. — *Aus* II. *folgt* III. Für jede Menge $Q \in \mathfrak{Q}$ ist nach (5) und (6) $\mu(Q) = \sup\big(I(g);\ 0 \leqq g \leqq \chi_Q,\ g \in \mathfrak{R}\big)$. Da Q relativ kompakt ist, besitzt jede Funktion $g \in \mathfrak{R}$ mit $0 \leqq g \leqq \chi_Q$ einen kompakten Träger und es ist somit $I(g) = 0$. Hieraus folgt $\mu(Q) = 0$. — *Aus* III. *folgt* I., da nach der Behauptung I des Satzes 10 $I_c = 0$ ist.

Dieses Korollar gestattet nunmehr die Herleitung einer einfachen Formel für die Berechnung von I_p und damit auch von I_c.

Satz 11. *Für jede Funktion $f \gtrsim 0$ aus $\mathfrak{R}$ ist*

$$I_p(f) = \lim_{\alpha \to +0} I\big(\min(\alpha, f)\big). \tag{26}$$

Das Funktional I ist dann und nur dann stetig bzw. rein-unstetig, wenn

$$\lim_{\alpha \to +0} I\big(\min(\alpha, f)\big) = 0 \tag{27}$$

bzw.

$$I\big(\min(1, f)\big) = I(f) \tag{28}$$

ist für alle $f \gtrsim 0$ aus $\mathfrak{R}$.

Beweis. Es sei f eine nicht negative Funktion aus $\mathfrak{R}$. Da f im Unendlichen verschwindet, ist die Funktion $g_\alpha = \min(\alpha, f) - f$ für jede reelle Zahl $\alpha > 0$ ein Element von $\mathfrak{R} \cap \mathfrak{R}(E)$. Nach dem Korollar zu Satz 10 gilt also $I_p(g_a) = 0$ und somit die Gleichung $I\big(\min(\alpha, f)\big) = I_c\big(\min(\alpha, f)\big) + I_p(f)$ für jedes $\alpha > 0$. Läßt man in dieser Gleichung α gegen $+0$ konvergieren, so folgt (26); wegen der Stetigkeit von I_c ist nämlich $\lim\limits_{\alpha \to +0} I_c\big(\min(\alpha, f)\big) = \lim\limits_{n \to \infty} I_c\left(\min\left(\frac{1}{n}, f\right)\right)$ $= I_c(0) = 0$. — Aus (26) folgt unmittelbar die Kennzeichnung der stetigen Funktionale durch Gl. (27). — Ist I rein-unstetig, so gilt nach dem Bisherigen $I(g_1) = 0$, also (28). Aus (28) folgt umgekehrt die Gleichung $I\big(\min(\alpha, f)\big) = I(f)$ für jedes $\alpha > 0$ und hieraus $I = I_p$ nach (26).

Wir erhalten als Korollar folgende Aussage über die Möglichkeit einer Integraldarstellung von $I(f)$.

Korollar. *Es sei f eine nicht negative Funktion aus $\mathfrak{R}$. Dann ist die Bedingung (27) notwendig und hinreichend für das Bestehen der Gleichung*

$$I(f) = \int f \, d\varrho = \overset{(u)}{\int} f \, d\mu .$$

Setzen wir $G = [f > 0]$, so ist für das Bestehen dieser Gleichungen dagegen nur hinreichend die Bedingung

$$\sup\big(I(g);\, 0 \leq g \leq \chi_G,\, g \in \mathfrak{R}\big) < +\infty \text{ [20]}.$$

Beweis. Daß die Bedingung (27) notwendig und hinreichend ist, folgt aus (25) und (26) sowie aus $I = I_\varepsilon + I_p$. — Es sei $\gamma(G) = \sup\big(I(g);\, 0 \leq g \leq \chi_G,\, g \in \mathfrak{R}\big)$. Wegen $0 \leq \frac{1}{\alpha} \min(\alpha, f) \leq \chi_G$ ist dann $0 \leq I\big(\min(\alpha, f)\big) \leq \alpha\gamma(G)$ für jedes $\alpha > 0$. Hieraus folgt die Bedingung (27).

3.6. *Beispiele rein-unstetiger, positiver, linearer Funktionale.*

1. Der lokal kompakte Raum E sei das halboffene Einheitsintervall $(0, 1] = \{x: 0 < x \leq 1\}$ der Zahlengeraden, versehen mit der üblichen Topologie. $\mathfrak{R}$ bestehe aus allen stetigen Funktionen $f|[0, 1]$, aufgefaßt als nur auf E definiert, mit $f(0) = 0$ und in $x = 0$ existierender rechtsseitiger Ableitung $f'(0)$. Für jedes $f \in \mathfrak{R}$ sei $I(f) = f'(0)$ [21].

2. E sei die Halbgerade $[0, +\infty) = \{x: 0 \leq x < +\infty\}$ mit der üblichen Topologie. $\mathfrak{R}$ bestehe aus allen stetigen Funktionen $f|[0, +\infty)$, für welche $\lim_{x \to +\infty} x f(x)$ (eigentlich) existiert. Für jedes $f \in \mathfrak{R}$ sei $I(f) = \lim_{x \to +\infty} x f(x)$.

Bemerkung. Im Beispiel 2 ist $I(f) = 0 = \int f \, d\varrho = \overset{(u)}{\int} f \, d\mu$ für die spezielle, nicht negative Funktion $f(x) = (1 + x^2)^{-1}$ $(x \in E)$ aus $\mathfrak{R}$; ferner ist $G = [f > 0]$ der Gesamtraum E. Bezeichnet q die Funktion $(1 + x)^{-1}$ aus $\mathfrak{R}$, so ist für jede reelle Zahl $\alpha > 0$ die Funktion $g_\alpha = \min(1, \alpha q)$ ein Element von $\mathfrak{R}$ und $0 \leq g_\alpha \leq 1$. Wegen $I(g_\alpha) = \alpha$ ist folglich $\sup\big(I(g);\, 0 \leq g \leq \chi_G,\, g \in \mathfrak{R}\big) = +\infty$. Die im Korollar zu Satz 11 genannte hinreichende Bedingung ist also nicht notwendig.

§ 4. Zurückführung des abstrakten Falles auf den konkreten

4.1. Im letzten Paragraphen ergab sich unter anderem, daß unter den dort gemachten speziellen Annahmen über E und $\mathfrak{R}$ zu jedem positiven, linearen Funktional $I|\mathfrak{R}$ ein Radonsches Maß ϱ auf E derart existiert, daß die Theorie des Riemann-Integrals bezüglich I mit der bezüglich ϱ identisch ist. Es bedeutet also keine Einschränkung der Allgemeinheit, im konkreten Fall eines lokal kompakten Raumes den Standpunkt von Haupt und Pauc einzunehmen. In den folgenden Betrachtungen wird dieses Ergebnis dahingehend verschärft, daß die Loomissche Theorie des Riemann-Integrals selbst

[20]) Vgl. auch Satz 16.

[21]) Nach Loomis [14], S. 174, wo das Beispiel aber nicht in unserem Sinne interpretiert wird.

im abstrakten Fall keinen wesentlichen Gewinn an Allgemeinheit gegenüber dem konkreten Fall bringt. Wir behaupten nämlich: Ist E eine Menge und $\mathfrak{R}$ ein System von beschränkten Funktionen $g|E$ mit den Eigenschaften (1) bis (3) sowie (21) und (22), so existiert ein lokal kompakter Raum $E'_{\mathfrak{R}}$, der E als dichte Teilmenge enthält, mit folgender Eigenschaft: Zu jedem positiven, linearen Funktional $I|\mathfrak{R}$ gibt es ein positives Radonsches Maß ϱ' auf $E'_{\mathfrak{R}}$ derart, daß eine Funktion $f|E$ dann und nur dann Riemann-integrierbar bezüglich I ist, wenn sie zu einer bezüglich ϱ' Riemann-integrierbaren Funktion $f'|E'_{\mathfrak{R}}$ fortgesetzt werden kann; für jede solche Fortsetzung f' von f ist $\int f\,d\mu = \int f'\,d\varrho'$.

4.2. Es liege der abstrakte Fall des § 1 vor: E sei eine Menge, $\mathfrak{R}$ ein System auf E definierter, reeller Funktionen mit den Eigenschaften (1) bis (3) und $I|\mathfrak{R}$ ein positives, lineares Funktional. Wir behaupten zunächst, daß es im Hinblick auf die Loomissche Theorie des Riemann-Integrals keine wesentliche Einschränkung bedeutet, über $\mathfrak{R}$ folgende drei Zusatz-Voraussetzungen zu machen: 1. *Jede Funktion aus* $\mathfrak{R}$ *ist beschränkt*; 2. *es besitzt* $\mathfrak{R}$ *die Eigenschaft* (21); 3. *es besitzt* $\mathfrak{R}$ *die Eigenschaft* (22).

Zu 1. Die Teilmenge $\mathfrak{R}^*$ von $\mathfrak{R}$ aller beschränkten Funktionen aus $\mathfrak{R}$ genügt den Bedingungen (1) bis (3). Bezeichnet $I^*|\mathfrak{R}^*$ die Verengerung von I auf $\mathfrak{R}$, so führen die Funktionale $I|\mathfrak{R}$ und $I^*|\mathfrak{R}^*$ zum gleichen System $\mathfrak{Q}$ quadrierbarer Mengen, zum gleichen Inhalt $\mu|\mathfrak{Q}$, zum gleichen System $\mathfrak{S}$ Riemann-integrierbarer Funktionen und zum gleichen Riemann-Integral über $\mathfrak{S}$. Zum Beweis sei Q eine I-quadrierbare Menge und $\varepsilon > 0$ eine reelle Zahl. Dann existieren Funktionen $g, h \in \mathfrak{R}$ mit $0 \leq g \leq \chi_Q \leq h$ und $I(h-g) \leq \varepsilon$. Setzen wir $h^* = \min(1, h)$, so sind g und h^* Funktionen aus $\mathfrak{R}^*$ mit $0 \leq g \leq \chi_Q \leq h^* \leq h$ und $I^*(h^*-g) \leq I(h-g) \leq \varepsilon$. Also ist Q auch I^*-quadrierbar. Trivial ist, daß umgekehrt eine I^*-quadrierbare Menge auch I-quadrierbar ist. Da in die Definition des Inhalts $\mu|\mathfrak{Q}$ nur Funktionen aus $\mathfrak{R}^*$ eingehen, führen I und I^* zum gleichen Inhalt μ. Hieraus folgt aber der Rest der Behauptung.

Zu 2. Es bezeichne E_0 die Menge aller Elemente $x \in E$ mit $g(x) = 0$ für alle $g \in \mathfrak{R}$. Aus der Definition der Vervollständigung $\overline{\mathfrak{R}}$ von $\mathfrak{R}$ bezüglich I folgt, daß jede Funktion $f \in \overline{\mathfrak{R}}$ auf E_0 gleich Null ist. Hieraus folgt insbesondere, daß jede quadrierbare Menge Q zu E_0 fremd und jede Riemann-integrierbare Funktion auf E_0 gleich Null ist. Es genügt daher, die Funktionen $g \in \mathfrak{R}$ nur auf dem verengerten Definitionsbereich $E - E_0$ zu betrachten. (Den Trivialfall, in welchem $E - E_0$ die leere Menge ist, also $\mathfrak{R}$ nur aus der konstanten Funktion $g = 0$ besteht, können wir unberücksichtigt lassen.)

Zu 3. Es bezeichne R die folgende Äquivalenzrelation in E: Zwei Elemente $x, y \in E$ heißen äquivalent, wenn $g(x) = g(y)$ ist für jede Funktion $g \in \mathfrak{R}$. Es sei $E^* = E/R$ die Menge der Äquivalenzklassen bezüglich R; ohne Beschränkung der Allgemeinheit können wir E^* als eine Teilmenge von E annehmen, die aus jeder Äquivalenzklasse genau einen Punkt als Repräsentanten enthält. Das System $\mathfrak{R}^*$ der auf E^* verengerten Funktionen aus $\mathfrak{R}$

besitzt neben den Eigenschaften (1) bis (3) auch die Eigenschaft (22) auf E^*. Wir übertragen das Funktional I auf $\mathfrak{R}^*$ und erhalten ein positives, lineares Funktional $I^*|\mathfrak{R}^*$. Dann gilt: Eine Funktion $f^*|E^*$ liegt dann und nur dann in der Vervollständigung von $\mathfrak{R}^*$ bezüglich I^*, wenn sie zu einer in der Vervollständigung $\overline{\mathfrak{R}}$ von $\mathfrak{R}$ bezüglich I gelegenen Funktion $f|E$ fortgesetzt werden kann. Die Fortsetzung $f\in\overline{\mathfrak{R}}$ von f^* kann so gewählt werden, daß f auf jeder Äquivalenzklasse bezüglich R konstant ist. Hieraus folgt, daß eine Menge $Q^*\subseteq E^*$ dann und nur dann I^*-quadrierbar ist, wenn eine I-quadrierbare Menge $Q\subseteq E$ mit $Q^*=Q\cap E$ existiert; bezeichnet μ^* den sich aus I^* ableitenden Inhalt, so ist $\mu^*(Q^*)=\mu(Q)$. Es folgt weiter, daß eine Funktion $f^*|E^*$ dann und nur dann Riemann-integrierbar bezüglich I^* ist, wenn sie zu einer bezüglich I Riemann-integrierbaren Funktion $f|E$ fortgesetzt werden kann; es gilt dann $\int f^*\,d\mu^*=\int f\,d\mu$.

4.3. Von nun an genüge $\mathfrak{R}$ neben (1) bis (3) auch noch den in 4.2 genannten drei Zusatz-Voraussetzungen. Dann behaupten wir:

Satz 12. *Zu E und $\mathfrak{R}$ existiert ein lokal kompakter Raum $E'_{\mathfrak{R}}$ mit folgenden Eigenschaften:*

(29) *E ist eine dichte Teilmenge von $E'_{\mathfrak{R}}$;*

(30) *jede Funktion $g\in\mathfrak{R}$ kann (auf genau eine Weise) zu einer auf $E'_{\mathfrak{R}}$ stetigen, im Unendlichen verschwindenden Funktion g' fortgesetzt werden;*

(31) *die Menge $\mathfrak{R}'$ der stetigen Fortsetzungen g' aller Funktionen $g\in\mathfrak{R}$ besitzt auf $E'_{\mathfrak{R}}$ die Trennungseigenschaften* (21) *und* (22)[22]).

Beweis. Nach dem Lemma 1 ist die Menge $\mathfrak{A}_0$ aller (beschränkten) Funktionen auf E, die durch Funktionen aus $\mathfrak{R}$ gleichmäßig approximiert werden können, eine hinsichtlich gleichmäßiger Konvergenz abgeschlossene Algebra. Die kleinste, $\mathfrak{A}_0$ umfassende Algebra $\mathfrak{A}$, welche auch die auf E konstanten Funktionen enthält, besteht dann genau aus den Funktionen $f=\alpha+f_0$, wobei α eine reelle Zahl und $f_0\in\mathfrak{A}_0$ ist. Wir behaupten, daß auch $\mathfrak{A}$ hinsichtlich gleichmäßiger Konvergenz abgeschlossen ist. Wenn $\mathfrak{A}=\mathfrak{A}_0$ ist, also $\mathfrak{A}_0$ die auf E konstanten Funktionen enthält, ist dies trivial. Wir können daher annehmen, daß $f=0$ die einzige konstante Funktion in $\mathfrak{A}_0$ ist. Es sei $\{f_n\}$ eine gegen eine Funktion $f|E$ gleichmäßig konvergente Folge von Funktionen aus $\mathfrak{A}$; jedes f_n kann in der Form $f_n=\alpha_n+f_n^0$ mit $\alpha_n\in\mathsf{P}$ und $f_n^0\in\mathfrak{A}_0$ dargestellt werden. Ist α ein Limeswert der Folge $\{\alpha_n\}$, so existiert eine gegen α konvergente Teilfolge $\{\alpha_{n_\nu}\}$; folglich konvergiert $\{f_{n_\nu}^0\}$ gleichmäßig gegen $f-\alpha$, d.h. es ist $f-\alpha\in\mathfrak{A}_0$. Für einen zweiten Limeswert β von $\{\alpha_n\}$ folgt entsprechend $f-\beta\in\mathfrak{A}_0$. Also sind α,β endlich und die konstante Funktion $\alpha-\beta=(f-\beta)-(f-\alpha)$ liegt in $\mathfrak{A}_0$; gemäß unserer Annahme über $\mathfrak{A}_0$ ist also $\alpha=\beta$, d.h. es existiert der Grenzwert der Folge $\{\alpha_n\}$ und zwar ist $\alpha=\beta=\lim\alpha_n$. Dann konvergiert aber auch die Folge $\{f_n^0\}$ gleichmäßig auf E gegen eine Grenzfunktion $f_0\in\mathfrak{A}_0$ und es ist $f=\alpha+f_0\in\mathfrak{A}$. Auf Grund der festgestellten

[22]) In (21) und (22) sind also E und $\mathfrak{R}$ sinngemäß durch $E'_{\mathfrak{R}}$ und $\mathfrak{R}'$ zu ersetzen.

Eigenschaften von $\mathfrak{A}$ existiert nach GELFAND und ŠILOV [9][23]) ein kompakter Raum E'', welcher E als dichte Teilmenge enthält und auf welchen die Funktionen aus $\mathfrak{A}$ und nur diese stetig fortgesetzt werden können. Es bezeichne $\mathfrak{R}''$ bzw. $\mathfrak{A}_0''$ bzw. $\mathfrak{A}''$ die Menge der auf E'' stetig fortgesetzten Funktionen aus $\mathfrak{R}$ bzw. $\mathfrak{A}_0$ bzw. $\mathfrak{A}$; dann ist $\mathfrak{A}''$ die Menge aller stetigen Funktionen auf E''. Nun sind zwei Fälle zu unterscheiden. Erstens. Es sei $\mathfrak{A}=\mathfrak{A}_0$, d.h. es enthalte $\mathfrak{A}_0$ die auf E konstanten Funktionen. Dann ist E'' selbst ein Raum $E_{\mathfrak{R}}'$. Die Trennungseigenschaften (21) und (22) sind auf E'' erfüllt, da jede auf E'' stetige, reelle Funktion durch Funktionen aus $\mathfrak{R}''$ zunächst auf E und wegen der Dichtigkeit von E in E'' dann auf ganz E'' gleichmäßig approximiert werden kann. Zweitens. Es sei $\mathfrak{A}\neq\mathfrak{A}_0$, d.h. $\mathfrak{A}_0$ enthalte $f=0$ als einzige auf E konstante Funktion. In diesem Falle ist $\mathfrak{A}_0''$ ein maximales Ideal in $\mathfrak{A}''$. Nach STONE [20][24]) existiert folglich ein eindeutig bestimmter Punkt p_0'' in E'' derart, daß $\mathfrak{A}_0''$ die Menge aller stetigen Funktionen $f''|E''$ mit $f''(p_0'')=0$ ist. Der Unterraum $E'=E''-(p_0'')$ von E'' ist dann ein Raum $E_{\mathfrak{R}}'$. Die Trennungseigenschaften (21) und (22) ergeben sich auf Grund der gleichen Bemerkung wie im bereits erledigten Falle. Damit ist der Satz 12 vollständig bewiesen[25]).

Bemerkung. Aus dem Beweis des Satzes 12 (und wegen Satz 13 natürlich auch aus den Eigenschaften (29) bis (31) von $E_{\mathfrak{R}}'$) folgt, daß $E_{\mathfrak{R}}'$ *dann und nur dann kompakt* ist, wenn die konstante Funktion $g=1$ durch Funktionen aus $\mathfrak{R}$ gleichmäßig approximiert werden kann, also wegen (3) in $\mathfrak{R}$ liegt.

Der Vollständigkeit halber beweisen wir

Satz 13. *Der Raum $E_{\mathfrak{R}}'$ ist durch die Eigenschaften* (29) *bis* (31) *bis auf Homöomorphien, welche E punktweise festlassen, eindeutig bestimmt.*

Beweis. Es seien E_1' und E_2' zwei Räume $E_{\mathfrak{R}}'$. Wir setzen $E_i''=E_i'$, falls E_i' kompakt ist; andernfalls sei E_i'' die durch Adjunktion des unendlich fernen Punktes w_i entstehende Alexandroffsche Kompaktifizierung von E_i'. In jedem Falle ist E_i'' ein kompakter Raum, in dem E dicht liegt und auf den jede stetige, im Unendlichen verschwindende Funktion $f'|E_i'$ stetig fortgesetzt werden kann; insbesondere können die Funktionen aus $\mathfrak{R}$ auf E_i'' stetig fortgesetzt werden. Aus den Eigenschaften von E_i' und aus Lemma 4 folgt dann, daß jede stetige Funktion $f''|E_i''$ eine Darstellung $f''=\alpha+f_0''$ besitzt, wobei α eine reelle Zahl ist und f_0'' durch auf E_i'' stetig fortgesetzte Funktionen aus $\mathfrak{R}$ gleichmäßig approximiert werden kann ($i=1, 2$). Daher kann jede beschränkte Funktion $f|E$, die zu einer auf E_1'' stetigen Funktion fortgesetzt werden kann, auch zu einer auf E_2'' stetigen Funktion fortgesetzt werden und umgekehrt. Nach STONE [20][26]) folgt hieraus die Existenz einer Homöomorphie ψ von E_1'' auf E_2'', welche E punktweise festläßt. Da E in E_1'' und E_2'' dicht liegt, gilt

[23]) Vgl. hierzu auch NÖBELING und BAUER [16], S. 34 und 35.

[24]) Vgl. hierzu S. 25 der in [23]) zitierten Arbeit.

[25]) Einen anderen, nur einfache topologische Hilfsmittel verwendenden Beweis hat Verf. in [17], S. 65–67 gegeben.

[26]) Vgl. auch [16], S. 28.

dann $f_1''=f_2''\circ\psi$ für jede Funktion $f|E$, die zu einer stetigen Funktion $f_1''|E_1''$ und zu einer stetigen Funktion $f_2''|E_2''$ fortgesetzt werden kann. Ist E_1' nicht kompakt, so folgt hieraus $0=f_1''(w_1)=f_2''(\psi(w_1))$ für jedes $f\in\mathfrak{R}$; wegen (31) ist also auch E_2' nicht kompakt, vielmehr gilt $w_2=\psi(w_1)$. Durch Vertauschung der Rollen von E_1' und E_2' folgt, daß umgekehrt mit E_1' auch E_2' kompakt ist. Damit ist gezeigt, daß ψ in jedem Falle E_1' auf E_2' homöomorph abbildet. Auch die auf E_1' verengerte Homöomorphie ψ läßt E punktweise fest.

Der Satz 13 berechtigt uns, fernerhin von *dem* Raum $E_{\mathfrak{R}}'$ zu sprechen. Ist (wie in § 3) E selbst ein lokal kompakter Raum und $\mathfrak{R}$ ein System stetiger, im Unendlichen verschwindender Funktionen auf E mit den Eigenschaften (1) bis (3) sowie (21) und (22), so ist $E_{\mathfrak{R}}'=E$.

Die aus Nr. 1.5 bekannte Bedingung (8) für $\mathfrak{R}$ erfährt jetzt eine einfache und neue Deutung.

Satz 14. *Die Funktionenmenge $\mathfrak{R}$ genügt (unter den bisherigen Annahmen) dann und nur dann der Bedingung* (8), *wenn die stetige Fortsetzung $g'|E_{\mathfrak{R}}'$ einer jeden Funktion $g\in\mathfrak{R}$ einen kompakten Träger besitzt.*

Beweis. Es genüge $\mathfrak{R}$ der Bedingung (8). Dann existiert zu beliebigem $g\in\mathfrak{R}$ eine nicht negative Funktion $h\in\mathfrak{R}$ derart, daß für jedes $x\in E$ aus $h(x)<1$ folgt $g(x)=0$. Bezeichnet g' bzw. h' die stetige Fortsetzung von g bzw. h auf $E_{\mathfrak{R}}'$, so folgt, da E in $E_{\mathfrak{R}}'$ dicht liegt, sogar für jedes $x'\in E_{\mathfrak{R}}'$ aus $h'(x')<1$ stets $g'(x')=0$. Da $h'\geqq 0$ ist und im Unendlichen verschwindet, ist $[h'\geqq 1]$ eine kompakte Menge. g' ist im Komplement dieser Menge gleich Null und besitzt daher einen kompakten Träger. — Umgekehrt sei der Träger $T(g')$ der stetigen Fortsetzung $g'|E_{\mathfrak{R}}'$ einer Funktion $g\in\mathfrak{R}$ kompakt. Dann gibt es eine nicht negative, stetige, im Unendlichen verschwindende Funktion $t'|E_{\mathfrak{R}}'$ mit $t'(x')=1$ für alle $x'\in T(g')$. Wegen (31) und Lemma 4 existiert zu t' eine nicht negative Funktion $h\in\mathfrak{R}$ derart, daß für deren stetige Fortsetzung $h'|E_{\mathfrak{R}}'$ gilt $0\leqq\frac{1}{2}h'\leqq t'\leqq\frac{1}{2}h'+\frac{1}{2}$. Diese Funktion h leistet bezüglich g das in (8) Verlangte: Aus $h(x)<1$ für ein $x\in E$ folgt $t'(x)<1$; also ist $x\notin T(g')$ und somit $g(x)=g'(x)=0$. Die Bedingung (8) ist also erfüllt.

4.4. Der Raum $E_{\mathfrak{R}}'$ wurde unabhängig von dem auf $\mathfrak{R}$ definierten Funktional I definiert. Erst jetzt beziehen wir I in unsere Betrachtungen mit ein. Für eine beliebige Funktion $g\in\mathfrak{R}$ bezeichne im folgenden g' deren eindeutig bestimmte stetige Fortsetzung auf $E_{\mathfrak{R}}'$; $\mathfrak{R}'$ sei die Menge der Fortsetzungen g' aller Funktionen $g\in\mathfrak{R}$. Wir übertragen I auf $\mathfrak{R}'$ durch die Festsetzung

$$I'(g')=I(g)\qquad(g\in\mathfrak{R}).\tag{32}$$

Dann ist auch $I'|\mathfrak{R}'$ ein positives, lineares Funktional. Da E in $E_{\mathfrak{R}}'$ dicht liegt, genügt auch $\mathfrak{R}'$ den Bedingungen (1) bis (3) für $\mathfrak{R}$; nach Satz 12 besitzt daher $\mathfrak{R}'$ bezüglich $E_{\mathfrak{R}}'$ alle Eigenschaften, die in § 3 $\mathfrak{R}$ bezüglich E hatte. Also sind die Resultate des § 3 auf $E_{\mathfrak{R}}'$, $\mathfrak{R}'$ und I' an Stelle von E, $\mathfrak{R}$ und I anwendbar. Es bezeichne $\mu'|\mathfrak{L}'$ den sich aus I' ableitenden, an $E_{\mathfrak{R}}'$ adaptierten Inhalt und ϱ' das eindeutig bestimmte positive Radonsche Maß auf $E_{\mathfrak{R}}'$, in

bezug auf welches alle Mengen $Q' \in \mathfrak{Q}'$ ϱ'-integrierbar mit $\varrho'(Q') = \mu'(Q')$ sind. Dann ist $\mathfrak{Q}'$ nach dem Korollar zu Satz 8 auch das System der ϱ'-quadrierbaren Mengen. Nach dem gleichen Korollar ist die Menge $\mathfrak{S}'$ der bezüglich I' Riemann-integrierbaren Funktionen gleich der Menge der bezüglich ϱ' Riemann-integrierbaren Funktionen. Es gilt nun der einleitend bereits angekündigte

Satz 15. I. *Eine Menge $Q \subseteq E$ ist dann und nur dann I-quadrierbar, wenn eine ϱ'-quadrierbare Menge $Q' \subseteq E'_{\mathfrak{R}}$ existiert mit $Q = Q' \cap E$. Für je zwei Mengen $Q \in \mathfrak{Q}$ und $Q' \in \mathfrak{Q}'$ mit $Q = Q' \cap E$ ist $\mu(Q) = \mu'(Q') = \varrho'(Q')$.*

II. *Eine reelle Funktion $f|E$ ist dann und nur dann Riemann-integrierbar bezüglich I, wenn sie zu einer bezüglich ϱ' Riemann-integrierbaren Funktion $f'|E'_{\mathfrak{R}}$ fortgesetzt werden kann. Für je zwei Funktionen $f \in \mathfrak{S}$ und $f' \in \mathfrak{S}'$, von denen die zweite eine Fortsetzung der ersten ist, gilt $\int f\, d\mu = \int f'\, d\mu' = \int f'\, d\varrho'$.*

Beweis. *Zu* I. Es sei $Q' \in \mathfrak{Q}'$; wir behaupten die I-Quadrierbarkeit der Menge $Q = Q' \cap E$ und die Gültigkeit der Gleichung $\mu(Q) = \mu'(Q')$ $(= \varrho'(Q'))$. Zu beliebigem $\varepsilon > 0$ existieren Funktionen $g', h' \in \mathfrak{R}'$ mit $g' \leq \chi'_{Q'} \leq h'$ [27]) und $I'(h' - g') \leq \varepsilon$. Für die zu $\mathfrak{R}$ gehörigen Verengerungen g und h von g' und h' auf E ist dann $g \leq \chi_Q \leq h$ und $I(h - g) \leq \varepsilon$. Also ist $Q \in \mathfrak{Q}$; ferner gilt $\mu(Q) = \mu'(Q')$ nach (6) und (32). — Umgekehrt sei $Q \subseteq E$ I-quadrierbar; zu zeigen ist die Existenz einer Menge $Q' \in \mathfrak{Q}'$ mit $Q = Q' \cap E$. Für Funktionen $g, h \in \mathfrak{R}$ folgt aus $g \leq \chi_Q \leq h$ stets $g' \leq \chi'_{\overline{Q}} \leq h'$ für die stetigen Fortsetzungen g' und h' von g und h auf $E'_{\mathfrak{R}}$ und die abgeschlossene Hülle $\overline{Q}$ von Q in $E'_{\mathfrak{R}}$. Setzen wir $Q' = Q \cup (\overline{Q} - E)$, so gilt auch $g' \leq \chi'_{Q'} \leq h'$. Die Menge Q' leistet daher das Verlangte.

Zu II. Es sei f' eine Funktion aus $\mathfrak{S}'$ und f die Verengerung von f' auf E. Zu $\varepsilon > 0$ existieren dann Mengen $Q'_1, \ldots, Q'_n \in \mathfrak{Q}'$ und reelle Zahlen $\alpha_1, \beta_1, \ldots, \alpha_n, \beta_n$ derart, daß $g' = \sum_{\nu=1}^{n} \alpha_\nu \chi'_{Q'_\nu} \leq f' \leq \sum_{\nu=1}^{n} \beta_\nu \chi'_{Q'_\nu} = h'$ und $\sum_{\nu=1}^{n} (\beta_\nu - \alpha_\nu) \mu'(Q'_\nu) \leq \varepsilon$ ist. Für die auf E definierten Funktionen $g = \sum \alpha_\nu \chi_{Q_\nu}$ und $h = \sum \beta_\nu \chi_{Q_\nu}$ mit $Q_\nu = Q'_\nu \cap E$, $\nu = 1, \ldots, n$, gilt folglich $g \leq f \leq h$ sowie $\sum (\beta_\nu - \alpha_\nu) \mu(Q_\nu) \leq \varepsilon$. Also ist $f \in \mathfrak{S}$. Da $\mu(Q_\nu) = \mu'(Q'_\nu)$ ist folgt ferner noch $\int f\, d\mu = \int f'\, d\mu'$ $(= \int f'\, d\varrho')$. Umgekehrt sei f eine Funktion aus $\mathfrak{S}$; wir beweisen die Existenz einer in $\mathfrak{S}'$ gelegenen Fortsetzung f' von f auf $E'_{\mathfrak{R}}$. Da jede Funktion aus $\mathfrak{S}$ gleich der Differenz zweier nicht negativer Funktionen aus $\mathfrak{S}$ ist, können wir ohne Beschränkung der Allgemeinheit $f \geq 0$ annehmen. Wie in 1.4 bezeichne $\mathfrak{E}$ die Menge aller $\mathfrak{Q}$-Treppenfunktionen; entsprechend sei $\mathfrak{E}'$ die Menge aller $\mathfrak{Q}'$-Treppenfunktionen (auf E'). Zu f existieren zwei Folgen $\{g_n\}_{n=1,2,\ldots}$ und $\{h_n\}_{n=1,2,\ldots}$ von Funktionen aus $\mathfrak{E}$ derart, daß $0 \leq g_1 \leq \cdots \leq g_n \leq g_{n+1} \leq \cdots \leq f \leq \cdots \leq h_{n+1} \leq h_n \leq \cdots \leq h_1$ und $\int (h_n - g_n)\, d\mu \leq 1/n$ für alle $n = 1, 2, \ldots$ gilt. Es genügt zu zeigen, daß jede Funktion g_n bzw. h_n zu einer Funktion g'_n bzw. h'_n aus $\mathfrak{E}'$ fortgesetzt werden kann, und zwar so, daß $g'_n \leq g'_{n+1} \leq h'_{n+1} \leq h'_n$

[27]) Im folgenden bezeichnet χ_A die auf E definierte, charakteristische Funktion einer Menge $A \subseteq E$ und χ'_A die auf $E'_{\mathfrak{R}}$ definierte, charakteristische Funktion einer Menge $A' \subseteq E'_{\mathfrak{R}}$.

und $\int (h'_n - g'_n)\, d\mu' = \int (h_n - g_n)\, d\mu$ für jedes n ist. Dann folgt nämlich $\sup g'_n(x') \leq \inf h'_n(x')$ für jeden Punkt $x' \in E'_{\mathfrak{R}}$. Um also zu einer in $\mathfrak{S}'$ gelegenen Fortsetzung f' von f auf $E'_{\mathfrak{R}}$ zu gelangen, braucht man für einen beliebigen Punkt $x' \in E'_{\mathfrak{R}} - E$ als Funktionswert $f'(x')$ nur eine reelle Zahl ξ mit $\sup g'_n(x') \leq \xi \leq \inf h'_n(x')$ zu wählen. Die Existenz der Fortsetzungen g'_n und h'_n ergibt sich folgendermaßen. Jeder Menge $Q \in \mathfrak{Q}$ ordnen wir die Menge $Q' = Q \cup (\overline{Q} - E)$ zu, wobei wieder $\overline{Q}$ die abgeschlossene Hülle von Q in $E'_{\mathfrak{R}}$ bedeutet; nach dem Beweis der Behauptung I ist $Q' \in \mathfrak{Q}'$, $Q' \cap E = Q$ und $\mu'(Q') = \mu(Q)$. Ist g eine beliebige nicht negative Funktion aus $\mathfrak{E}$ und $g = \sum \lambda_\nu \chi_{Q_\nu}$ eine ihrer Darstellungen mit paarweise fremden Mengen $Q_1, \ldots, Q_n \in \mathfrak{Q}$, so setzen wir sie zu einer nicht negativen Funktion $g' | E'_{\mathfrak{R}}$ wie folgt fort: Für einen Punkt $x' \in E'_{\mathfrak{R}} - E$ setzen wir $g'(x') = 0$, falls x' in keiner der Mengen $Q'_1, \ldots, Q'_n$ liegt; liegt dagegen x' in $Q'_{\nu_1}, \ldots, Q'_{\nu_r}$, $\nu_1 < \cdots < \nu_r$, aber in keiner weiteren Menge Q'_ν, so setzen wir $g'(x') = \max(\lambda_{\nu_1}, \ldots, \lambda_{\nu_r})$. Für den Beweis der Unabhängigkeit der so definierten Fortsetzung g' von der gewählten Darstellung von g seien $g = \sum_{\mu=1}^{m} \alpha_\mu \chi_{A_\mu}$ und $g = \sum_{\nu=1}^{n} \beta_\nu \chi_{B_\nu}$ zwei Darstellungen von g mit paarweise fremden Mengen $A_1, \ldots, A_m \in \mathfrak{Q}$ bzw. $B_1, \ldots, B_n \in \mathfrak{Q}$; ferner sei x' ein Punkt von $E'_{\mathfrak{R}} - E$. Wir beweisen zuerst folgende einfache Bemerkung. Liegt x' in B'_ν, ist $\beta_\nu \neq 0$ und sind $A_{\mu_1}, \ldots, A_{\mu_r}$ genau die zu B_ν nicht fremden Mengen unter den $A_1, \ldots, A_m$, so liegt x' in einer der Mengen $A'_{\mu_1}, \ldots, A'_{\mu_r}$ und es ist $\beta_\nu = \alpha_{\mu_1} = \cdots = \alpha_{\mu_r}$. Wegen $x' \in B'_\nu$ ist nämlich $B_\nu \neq \emptyset$ und wegen $\beta_\nu \neq 0$ existiert folglich wenigstens ein A_μ mit $A_\mu \cap B_\nu \neq \emptyset$; da g auf A_μ und B_ν konstant ist, so ist also $\alpha_\mu = \beta_\nu$ für jedes μ mit $A_\mu \cap B_\nu \neq \emptyset$. Um die Gleichheit der Funktionswerte $g'(x')$ nachzuweisen, die man ausgehend von der einen bzw. anderen Darstellung von g erhält, genügt es, die beiden folgenden Fälle zu betrachten. 1. Fall. Es liegt x' in keiner der Mengen $A'_1, \ldots, A'_m$. 2. Fall. Es liegt x' in den Mengen $A'_{\mu_1}, \ldots, A'_{\mu_k}$ und $B'_{\nu_1}, \ldots, B'_{\nu_l}$, aber in keiner weiteren Menge unter den $A'_1, \ldots, A'_m$ bzw. $B'_1, \ldots, B'_n$. Im ersten Falle folgt aus der vorangestellten Bemerkung, daß x' höchstens in Mengen B'_ν mit $\beta_\nu = 0$ liegt. Im zweiten Falle folgt aus der gleichen Bemerkung, daß jedes von Null verschiedene $\alpha_{\mu_1}, \ldots, \alpha_{\mu_k}$ unter den $\beta_{\nu_1}, \ldots, \beta_{\nu_l}$ und umgekehrt jedes von Null verschiedene $\beta_{\nu_1}, \ldots, \beta_{\nu_l}$ unter den $\alpha_{\mu_1}, \ldots, \alpha_{\mu_k}$ auftritt. Da g und damit alle Koeffizienten α_μ, β_ν, die zu nicht leeren Mengen Q_μ, Q_ν gehören, nicht negativ sind, führen also in beiden Fällen beide Darstellungen von g zum gleichen Funktionswert $g'(x')$. Weiter behaupten wir, daß g' eine Funktion aus $\mathfrak{E}'$ ist. Hierzu sei $g = \sum_{\nu=1}^{n} \lambda_\nu \chi_{Q_\nu}$ eine Darstellung von g; die Mengen $Q_\nu \in \mathfrak{Q}$ seien wieder paarweise fremd und es sei $0 < \lambda_1 < \cdots < \lambda_n$. Wir setzen $Q' = \bigcup_{\nu=1}^{n} Q'_\nu$ und $\tilde{Q}'_\nu = Q' - Q'_\nu$, $\nu = 1, \ldots, n$; dann sind Q' und $\tilde{Q}'_\nu$ Mengen aus $\mathfrak{Q}'$. Die Urbildmenge $U'_\nu = [g' = \lambda_\nu]$ ist definitionsgemäß für jedes $\nu = 1, \ldots, n$ gleich $Q'_\nu \cap \tilde{Q}'_{\nu+1} \cap \cdots \cap \tilde{Q}'_n$, also ein Element von $\mathfrak{Q}'$. Da g' nur die Werte $0, \lambda_1, \ldots, \lambda_n$ annehmen kann, ist $g' = \sum_{\nu=1}^{n} \lambda_\nu \chi'_{U'_\nu}$

und somit $g' \in \mathfrak{E}'$. Wegen $U_\nu' \cap E = Q_\nu$ gilt nach Behauptung I ferner $\mu'(U_\nu') = \mu(Q_\nu)$ für jedes ν und somit $\int g' d\mu' = \int g \, d\mu$. Schließlich zeigen wir, daß aus $0 \leq g \leq h$, $g, h \in \mathfrak{E}$, stets $g' \leq h'$ folgt. Hierzu wählen wir in Ω paarweise fremde Mengen $Q_1, \ldots, Q_n$, so daß $g = \sum_{\nu=1}^{n} \alpha_\nu \chi_{Q_\nu}$ und $h = \sum_{\nu=1}^{n} \beta_\nu \chi_{Q_\nu}$ ist. Dann gilt $\alpha_\nu \leq \beta_\nu$ für jedes ν mit $Q_\nu \neq \emptyset$ und aus der Definition von g' und h' folgt sofort $g' \leq h'$. Wendet man das beschriebene Verfahren der Fortsetzung nicht negativer Funktionen aus $\mathfrak{E}$ zu nicht negativen Funktionen aus $\mathfrak{E}'$ auf die Funktionen g_n und h_n an, so erhält man Fortsetzungen $g_n', h_n' \in \mathfrak{E}'$ mit den gewünschten Eigenschaften. Also ist nunmehr auch die Behauptung II vollständig bewiesen.

Korollar. *Für jede Funktion $f \in \mathfrak{R}$ ist* $\overset{(u)}{\int} f \, d\mu = \overset{(u)}{\int} f' \, d\mu' = \int f' \, d\varrho'$.

Beweis. Nach Satz 10 ist $\overset{(u)}{\int} f' \, d\mu' = \int f' \, d\varrho'$. Es genügt also die Gleichung $\overset{(u)}{\int} f \, d\mu = \overset{(u)}{\int} f' \, d\mu'$ und diese wiederum nur für nicht negative Funktionen $f \in \mathfrak{R}$ zu beweisen. Für ein $f \geq 0$ aus $\mathfrak{R}$ sei g eine Funktion aus $\mathfrak{S}$ mit $0 \leq g \leq f$. Ist $g' \in \mathfrak{S}'$ eine Fortsetzung von g auf $E'_{\mathfrak{R}}$, so ist deren Positivteil $g'^{+} \in \mathfrak{S}'$ und eine nicht negative Fortsetzung von g auf $E'_{\mathfrak{R}}$; nach (11) ist dann $g'^{*} = \min(g'^{+}, f')$ ein Element von $\mathfrak{S}'$ sowie eine Fortsetzung von g auf $E'_{\mathfrak{R}}$ mit $0 \leq g'^{*} \leq f'$. Nach Satz 15 ist daher die Menge der auf E verengerten Funktionen $g' \in \mathfrak{S}'$ mit $0 \leq g' \leq f'$ identisch mit der Menge aller Funktionen $g \in \mathfrak{S}$ mit $0 \leq g \leq f$. Aus der Definition des uneigentlichen Riemann-Integrals und Satz 15 folgt daher $\overset{(u)}{\int} f \, d\mu = \overset{(u)}{\int} f' \, d\mu'$.

4.5. Das in § 3 bewiesene Korollar zu Satz 11 über die Darstellbarkeit von $I(f)$ durch ein Integral läßt sich jetzt wie folgt verallgemeinern.

Satz 16. *Für eine nicht negative Funktion $f \in \mathfrak{R}$ bestehen die Gleichungen*

$$I(f) = \overset{(u)}{\int} f \, d\mu = \overset{(u)}{\int} f' \, d\mu' = \int f' \, d\varrho' \tag{33}$$

dann und nur dann, wenn

$$\lim_{\alpha \to +0} I\big(\min(\alpha, f)\big) = 0 \tag{34}$$

ist. Die Bedingung (34) ist insbesondere dann erfüllt, wenn für die Menge G aller Elemente $x \in E$ mit $f(x) > 0$

$$\sup\big(I(g); 0 \leq g \leq \chi_G, g \in \mathfrak{R}\big) < +\infty \tag{35}$$

ist[28]).

Beweis. Da $\min(\alpha, f')$ die Funktion $\min(\alpha, f)$ auf $E'_{\mathfrak{R}}$ stetig fortsetzt, folgt die Gleichwertigkeit von (33) und (34) aus den Korollaren zu den

[28]) Ist f eine beliebige Funktion aus $\mathfrak{R}$, so ist die gleiche Bedingung (35) hinreichend für die Gültigkeit der Gln. (33), wenn hierbei G als Menge aller $x \in E$ mit $|f(x)| > 0$ definiert wird.

Sätzen 11 und 15. Daß aus (35) die Gl. (34) folgt, wurde bereits im Beweis des Korollars zu Satz 11 gezeigt.

Bemerkungen. 1. Durch Satz 16 haben wir auf anderem Wege ein Resultat von Loomis [*14*], S. 176 gewonnen. Danach ist im abstrakten Fall des § 1 für eine beschränkte, nicht negative Funktion $f \in \mathfrak{R}$ die Bedingung (34) notwendig und hinreichend für das Bestehen der Gleichung $I(f) = \overset{(u)}{\int} f\, d\mu$.

2. Aus dem Beweis des Satzes 16 und des Korollars zu Satz 11 kann folgende Deutung der Bedingung (34) entnommen werden. Sie ist gleichbedeutend damit, daß der rein-unstetige Teil des Funktionals $I' | \mathfrak{R}'$ für die stetige Fortsetzung f' von f auf $E'_{\mathfrak{R}}$ gleich Null ist.

3. Die Bedingung (35) tritt in ähnlichem Zusammenhang auch bei Hewitt [*13*], S. 275 auf. Dort wird wie bei uns von einem positiven, linearen Funktional I über einem (dort mit $\mathfrak{F}$ bezeichneten) System $\mathfrak{R}$ von auf einer Menge E definierten, reellen Funktionen ausgegangen; von $\mathfrak{R}$ werden die Eigenschaften (1) bis (3) gefordert. Es wird das Ziel verfolgt, eine Darstellung von $I(f)$ durch ein Unterteilungsintegral zu finden, dem eine nach dem Vorbild der Theorie Radonscher Maße konstruierte, additive Mengenfunktion γ^* zugrunde liegt. In Theorem 2.27 wird die Gültigkeit der Gleichung $I(f) = \int f d\gamma^*$ für jede *beschränkte* Funktion $f \in \mathfrak{R}$ behauptet, welche der Bedingung (35) mit $G = [|f| > 0]$ genügt. Nach einer brieflichen Mitteilung von Herrn Hewitt ist sein Darstellungssatz (jedenfalls in dieser Allgemeinheit) falsch. Unser Satz 16 zeigt, daß das Hewittsche Ergebnis jedenfalls dann richtig ist, wenn man $\int f d\gamma^*$ entweder durch $\overset{(u)}{\int} f\, d\mu$ oder durch $\int f' d\varrho'$ ersetzt. (Für die zweite Ersetzung müssen allerdings an $\mathfrak{R}$ die in 4.2 formulierten drei Zusatzforderungen gestellt werden, die aber letzten Endes keine Beschränkung der Allgemeinheit bedeuten.) Herr Hewitt hat übrigens selbst angeregt, eine Richtigstellung seines Satzes entweder durch Verwendung eines anderen Integralbegriffes oder durch geeignete Einführung idealer Punkte zu versuchen.

4.6. Abschließend wenden wir uns der Frage zu, in welcher Weise sich der stetige Teil I_c des Funktionals I durch das auf $E'_{\mathfrak{R}}$ definierte Radonsche Maß ϱ' ausdrückt.

Satz 17. *Es sei f eine nicht negative Funktion aus $\mathfrak{R}$ und $\tilde{f}$ diejenige Fortsetzung von f auf $E'_{\mathfrak{R}}$, welche in allen Punkten von $E'_{\mathfrak{R}} - E$ gleich Null ist. Dann gilt für den stetigen Teil I_c von I*

$$I_c(f) = \overline{\int} \tilde{f} d\varrho' \text{ [29]}. \tag{36}$$

Beweis. Zunächst sei an die Definition von $\overline{\int} \tilde{f} d\varrho' = N(\tilde{f})$ erinnert. Es ist

$$\overline{\int} \tilde{f} d\varrho' = \inf\left(\sup \int k'_n d\varrho' ; \{k'_n\} \in \mathfrak{k}'(\tilde{f})\right),$$

wenn $\mathfrak{k}'(\tilde{f})$ die Menge aller nicht fallenden Folgen $\{k'_n\}$ nicht negativer Funktionen $k'_n \in \mathfrak{R}(E'_{\mathfrak{R}})$ mit $\tilde{f} \leq \sup k'_n$ bezeichnet. Ist andererseits $\mathfrak{r}(f)$ bzw. $\mathfrak{r}'(\tilde{f})$

[29]) Bezüglich $\overline{\int} \tilde{f} d\varrho'$ vgl. Fußnote [17]).

die Menge aller nicht fallenden Folgen $\{f_n\}$ bzw. $\{f'_n\}$ nicht negativer Funktionen $f_n \in \mathfrak{R}$ bzw. $f'_n \in \mathfrak{R}'$ mit $f \leq \sup f_n$ bzw. $\tilde{f} \leq \sup f'_n$, so kann (nach einer einfachen Überlegung) die in Lemma 3 gewonnene Formel für $I_c(f)$ auch so geschrieben werden

$$I_c(f) = \inf\left(\sup I(f_n); \{f_n\} \in \mathfrak{r}(f)\right).$$

Nun ist offenbar für eine Folge $\{f_n\} \in \mathfrak{r}(f)$ die Folge $\{f'_n\}$ der stetigen Fortsetzungen f'_n der Funktionen f_n auf $E'_{\mathfrak{R}}$ ein Element von $\mathfrak{r}'(\tilde{f})$ und umgekehrt für eine Folge $\{f'_n\} \in \mathfrak{r}'(\tilde{f})$ die Folge $\{f_n\}$ der Verengerungen f_n von f'_n auf E ein Element von $\mathfrak{r}(f)$. Wegen (32) kann daher die letzte Formel zu

$$I_c(f) = \inf\left(\sup I'(f'_n); \{f'_n\} \in \mathfrak{r}'(\tilde{f})\right)$$

umgeschrieben werden. Wir beweisen die Gleichheit der rechten Seite dieser Gleichung mit $\overline{\int} \tilde{f}\, d\varrho'$. Für jede Folge $\{f'_n\} \in \mathfrak{r}'(\tilde{f})$ gilt nach bekannten Eigenschaften des Oberintegrals $\overline{\int} \tilde{f}\, d\varrho' \leq \sup \overline{\int} f'_n\, d\varrho'$; da nach (25) $\overline{\int} f'_n\, d\varrho' = \int f'_n\, d\varrho' \leq I'(f'_n)$ ist und $\{f'_n\}$ in $\mathfrak{r}'(\tilde{f})$ beliebig gewählt war, folgt $\overline{\int} \tilde{f}\, d\varrho' \leq I_c(f)$. Ist umgekehrt $\{k'_n\}$ eine beliebige Folge aus $\mathfrak{k}'(\tilde{f})$, so ist die Verengerung k_n einer jeden Funktion k'_n auf E nach Satz 15 ein Element von $\mathfrak{E}$ und $\int k_n\, d\mu = \int k'_n\, d\varrho'$. Aus $0 \leq k_1 \leq \cdots \leq k_n \leq k_{n+1} \leq \cdots$ und $f \leq \sup k_n$ folgt $I_c(f) \leq \sup \overline{I}_c(k_n) \leq \sup \overline{I}(k_n)$ nach Satz 4, Beh. I; wegen $\overline{I}(k_n) = \int k_n\, d\mu = \int k'_n\, d\varrho'$ ist also $I_c(f) \leq \sup \int k'_n\, d\varrho'$ und somit $I_c(f) \leq \overline{\int} \tilde{f}\, d\varrho'$. Damit ist (36) bewiesen.

Nach 3.4 ist der sich aus $I' | \mathfrak{R}'$ ableitende, an $E'_{\mathfrak{R}}$ adaptierte Inhalt $\mu' | \mathfrak{Q}'$ σ-additiv. Aus μ' leitet sich daher das äußere Maß $\overline{\mu'}$ und das innere Maß $\underline{\mu'}$ ab; $\overline{\mu'}$ und $\underline{\mu'}$ sind definiert auf dem System aller Teilmengen von $E'_{\mathfrak{R}}$, welche durch eine Folge von Mengen aus $\mathfrak{Q}'$ überdeckt werden können[30]). Mit Hilfe von $\overline{\mu'}$ und $\underline{\mu'}$ errechnet sich der σ-additive Teil μ_c und der reinendlich-additive Teil μ_p von $\mu | \mathfrak{Q}$ wie folgt.

Satz 18. *Für jede Menge $Q \in \mathfrak{Q}$ ist*

$$\mu_c(Q) = \overline{\mu'}(Q) \quad \text{und} \quad \mu_p(Q) = \underline{\mu'}(\overline{Q} - Q), \tag{37}$$

wenn $\overline{Q}$ die abgeschlossene Hülle von Q in $E'_{\mathfrak{R}}$ bezeichnet.

Beweis. Zu $Q \in \mathfrak{Q}$ und $\varepsilon > 0$ existieren nach (15) Funktionen $g, h \in \mathfrak{R}$ mit $0 \leq g \leq \chi_Q \leq h$, $I_c(g) \leq \mu_c(Q) \leq I_c(h)$ und $I_c(h-g) \leq \varepsilon$. Bezeichnen $\tilde{g}$ und $\tilde{h}$ die in allen Punkten von $E'_{\mathfrak{R}} - E$ verschwindenden Fortsetzungen von g und h auf $E'_{\mathfrak{R}}$, so gilt $\tilde{g} \leq \chi'_Q \leq \tilde{h}$ und $\overline{\int} \tilde{g}\, d\varrho' \leq \overline{\int} \chi'_Q\, d\varrho' \leq \overline{\int} \tilde{h}\, d\varrho'$. Nach Satz 17 ist $\overline{\int} \tilde{g}\, d\varrho' = I_c(g)$ und $\overline{\int} \tilde{h}\, d\varrho' = I_c(h)$; hieraus folgt $|\overline{\int} \chi'_Q\, d\varrho' - \mu_c(Q)| \leq \varepsilon$. Da $\varepsilon > 0$ beliebig gewählt war, ist also $\mu_c(Q) = \overline{\int} \chi'_Q\, d\varrho'$. Der erste Teil der Behauptung folgt daher aus der Gleichung $\overline{\mu'}(Q) = \overline{\int} \chi'_Q\, d\varrho'$, deren Richtigkeit aus der Stoneschen Integrationstheorie entnommen werden kann. Der Vollständigkeit halber geben wir hierfür jedoch einen Beweis an. Es bezeichne $\mathfrak{m}'(Q)$ das System aller nicht fallenden Folgen $\{Q'_n\}$ von Mengen aus $\mathfrak{Q}'$ mit

[30]) Vgl. etwa HAUPT-AUMANN-PAUC [12b], S. 55.

$Q \subseteq \bigcup_n Q'_n$; nach Satz 15 ist $\mathfrak{m}'(Q)$ nicht leer. Nach Definition des äußeren Maßes ist

$$\overline{\mu'}(Q) = \inf\left(\sup \mu'(Q'_n) ; \{Q'_n\} \in \mathfrak{m}'(Q)\right).$$

Für jede Folge $\{Q'_n\} \in \mathfrak{m}'(Q)$ gilt daher $\overline{\int} \chi'_Q \, d\varrho' \leq \sup \overline{\int} \chi'_{Q'_n} \, d\varrho' = \sup \int \chi'_{Q'_n} \, d\varrho' = \sup \mu'(Q'_n)$; also ist $\overline{\int} \chi'_Q \, d\varrho' \leq \overline{\mu'}(Q)$. Weiter existiert zu jeder reellen Zahl ε mit $0 < \varepsilon < 1$ eine nicht fallende Folge $\{f'_n\}$ nicht negativer Funktionen $f'_n \in \mathfrak{R}(E'_{\mathfrak{R}})$ mit $\chi'_Q \leq \sup f'_n$ und $\sup \int f'_n \, d\varrho' \leq \overline{\int} \chi'_Q \, d\varrho' + \varepsilon$. Nach Satz 8 ist jedes f'_n Riemann-integrierbar bezüglich I', also insbesondere I'-meßbar; folglich existiert eine Folge $\{\alpha_n\}$ reeller Zahlen mit $\alpha_1 > \alpha_2 > \cdots > \alpha_n > \cdots > 1 - \varepsilon = \eta^{-1}$ derart, daß $Q'_n = [f'_n > \alpha_n]$ ein Element von $\mathfrak{Q}'$ ist. Nach Konstruktion gilt $\{Q'_n\} \in \mathfrak{m}'(Q)$ und $\mu'(Q'_n) \leq \alpha_n^{-1} \int f'_n \, d\varrho' \leq \eta \int f'_n \, d\varrho'$; also ist $\overline{\mu'}(Q) \leq \sup \mu'(Q'_n) \leq \eta \sup \int f'_n \, d\varrho' \leq \frac{1}{1-\varepsilon} (\overline{\int} \chi'_Q \, d\varrho' + \varepsilon)$. Hieraus folgt die noch fehlende Ungleichung $\overline{\mu'}(Q) \leq \overline{\int} \chi'_Q \, d\varrho'$ und somit die Gleichung $\overline{\mu'}(Q) = \overline{\int} \chi'_Q \, d\varrho'$. – Nach dem Beweis des Satzes 15 ist die Menge $Q' = Q \cup (\overline{Q} - E)$ für jedes $Q \in \mathfrak{Q}$ ein Element von $\mathfrak{Q}'$; also ist auch $\overline{Q} = \overline{Q'} \in \mathfrak{Q}'$ und $\mu(Q) = \mu'(Q') = \mu'(\overline{Q})$ nach Satz 15 und der Adaption von μ' an $E'_{\mathfrak{R}}$. Wegen $\overline{Q} \in \mathfrak{Q}'$ gilt $\mu'(\overline{Q}) = \overline{\mu'}(Q) + \underline{\mu'}(\overline{Q} - Q)$. Hieraus und aus $\mu(Q) = \mu_c(Q) + \mu_p(Q) = \overline{\mu'}(Q) + \mu_p(Q)$ folgt schließlich $\mu_p(Q) = \underline{\mu'}(\overline{Q} - Q)$.

Die Formel (37) verallgemeinert die erste Hälfte von (25). Genügen nämlich E und $\mathfrak{R}$ den in 3.1 formulierten Bedingungen, so ist $E'_{\mathfrak{R}} = E$, $\tilde{f} = f$ und $\mu' = \mu$, ferner ist dann f ϱ-integrierbar.

Korollar. *Die auf E konstanten Funktionen seien Elemente von $\mathfrak{R}$. Dann sind folgende Aussagen gleichwertig:* (1a) *$\mu | \mathfrak{Q}$ ist σ-additiv*; (2a) *es ist $\underline{\mu'}(E'_{\mathfrak{R}} - E) = 0$*; (3a) *es ist $\mu(E) = \overline{\mu'}(E)$. Ferner sind gleichwertig die Aussagen:* (1b) *$\mu | \mathfrak{Q}$ ist rein-endlich-additiv*; (2b) *es ist $\overline{\mu'}(E) = 0$*; (3b) *es ist $\mu(E) = \underline{\mu'}(E'_{\mathfrak{R}} - E)$.*

Beweis. Die Behauptung folgt unmittelbar aus Satz 18. Man hat nur zu beachten, daß $E \in \mathfrak{Q}$ und $\overline{E} = E'_{\mathfrak{R}}$ ist.

Bemerkung. Überträgt man von $\mathfrak{R}$ auf $\mathfrak{R}'$ auch die Zerlegungsteile I_c und I_p von I durch die zu (32) analogen Festsetzungen $(I_c)'(g') = I_c(g)$ und $(I_p)'(g') = I_p(g)$ $(g \in \mathfrak{R})$, so spiegelt sich die Zerlegung $I = I_c + I_p$ in der Zerlegung $I' = (I_c)' + (I_p)'$ wider. Im allgemeinen gilt nicht $(I_c)' = (I')_c$ und $(I_p)' = (I')_p$, sondern nur $(I_c)'(g') \leq (I')_c(g')$ und $(I_p)'(g') \geq (I')_p(g')$ für alle $g' \geq 0$ aus $\mathfrak{R}'$. Hierzu genügt es zu bemerken, daß mit $I_c | \mathfrak{R}$ auch $(I_c)' | \mathfrak{R}'$ stetig ist. *Notwendig und hinreichend für das Bestehen der Gleichungen $(I_c)' = (I')_c$ und $(I_p)' = (I')_p$ ist vielmehr die σ-Additivität des Inhalts $\mu | \mathfrak{Q}$.*

Beweis. Für alle $f \in \mathfrak{R}$ ist $(I_c)'(f') = I_c(f) = \overset{(u)}{\int} f \, d\mu_c$ nach Satz 5 sowie $(I')_c(f') = \overset{(u)}{\int} f' \, d\mu' = \overset{(u)}{\int} f \, d\mu$ nach Satz 10 und dem Korollar zu Satz 15. Folglich gilt $(I_c)' = (I')_c$ und damit auch $(I_p)' = (I')_p$ dann und nur dann, wenn $\overset{(u)}{\int} f \, d\mu_c = \overset{(u)}{\int} f \, d\mu$ ist für alle $f \in \mathfrak{R}$. Nach Satz 6 ist hiermit $\mu_c = \mu$

gleichwertig. — Falls μ σ-additiv ist, ergibt sich aus Satz 10 als Darstellung für den stetigen Teil von I die für alle $f \in \mathfrak{R}$ gültige Gleichung $I_c(f) = \int f' \, d\varrho'$.

4.7. Beispiel. *Die Stonesche Kompaktifizierung einer Booleschen Mengenalgebra als Beispiel für einen Raum $E'_{\mathfrak{R}}$.* Es sei $\mathfrak{B}$ eine Boolesche Algebra von Teilmengen einer Grundmenge E, die in $\mathfrak{B}$ als Element enthalten ist; die Menge E sei durch $\mathfrak{B}$ separiert, d. h. zu je zwei verschiedenen Elementen $x, y \in E$ existiert ein $A \in \mathfrak{B}$ mit $x \in A$ und $y \in E - A$. Mit $\mathfrak{R}$ werde bezeichnet die Menge aller $\mathfrak{B}$-Treppenfunktionen, also aller Funktionen $g = \sum_{\nu=1}^{n} \lambda_\nu \chi_{A_\nu}$, wobei $\lambda_1, \ldots, \lambda_n$ reelle Zahlen und $A_1, \ldots, A_n$ Mengen aus $\mathfrak{B}$ sind. Dann besitzt $\mathfrak{R}$ die Eigenschaften (1) bis (3) sowie (21) und (22), ferner ist jede Funktion $g \in \mathfrak{R}$ beschränkt. Also sind die in 4.2 gemachten Voraussetzungen über $\mathfrak{R}$ erfüllt und es existiert der Raum $E'_{\mathfrak{R}}$. Dieser ist kompakt, da die konstante Funktion $g = 1$ in $\mathfrak{R}$ liegt.

Nach STONE [*20*] existiert zu jeder Booleschen Algebra $\mathfrak{B}$ ein kompakter, total unzusammenhängender Raum $U_{\mathfrak{B}}$ und ein Isomorphismus Φ von $\mathfrak{B}$ auf die Boolesche Mengenalgebra aller zugleich offenen und abgeschlossenen Teilmengen von $U_{\mathfrak{B}}$. Die Punkte von $U_{\mathfrak{B}}$ sind die Ultrafilter in $\mathfrak{B}$; für ein beliebiges Element $A \in \mathfrak{B}$ besteht ΦA aus allen Ultrafiltern in $\mathfrak{B}$, die A als Element enthalten. Die Mengen ΦA mit $A \in \mathfrak{B}$ bilden eine offene (und abgeschlossene) Basis von $U_{\mathfrak{B}}$. Wendet man das Stonesche Verfahren auf die Boolesche Mengenalgebra $\mathfrak{B}$ an, so kann E als dichte Teilmenge von $U_{\mathfrak{B}}$ aufgefaßt werden, indem man jeden Punkt $x \in E$ mit dem durch x in $\mathfrak{B}$ bestimmten Haupt(ultra)filter identifiziert. Für jedes $A \in \mathfrak{B}$ ist dann ΦA die abgeschlossene Hülle von A in $U_{\mathfrak{B}}$ und $A = E \cap \Phi A$[31]. Wir behaupten:

Es existiert eine die Punkte von E festlassende Homöomorphie von $E'_{\mathfrak{R}}$ auf $U_{\mathfrak{B}}$.

Beweis. Für jede Menge $A \in \mathfrak{B}$ ist $g = \chi_A$ als Element von $\mathfrak{R}$ zu einer auf $E'_{\mathfrak{R}}$ stetigen Funktion g' fortsetzbar. Aus der Stetigkeit von g' folgt $g'(x') = 1$ für alle $x' \in \bar{A}$ und $g'(y') = 0$ für alle $y' \in \overline{E - A}$; wegen $E'_{\mathfrak{R}} = \overline{E} = \bar{A} \cup \overline{E - A}$ ist daher $g' = \chi'_{\bar{A}}$. Hieraus folgt weiter, daß die Zuordnung $A \to \bar{A}$ ein Isomorphismus von $\mathfrak{B}$ auf die Boolesche Algebra $\overline{\mathfrak{B}}$ der abgeschlossenen Hüllen $\bar{A}$ von Mengen $A \in \mathfrak{B}$ ist und $A = E \cap \bar{A}$ für jede Menge $A \in \mathfrak{B}$ gilt. Die zugleich offenen und abgeschlossenen Mengen $\bar{A} \in \overline{\mathfrak{B}}$ bilden eine offene (und abgeschlossene) Basis von $E'_{\mathfrak{R}}$. Zu einem beliebigen Punkt $x' \in E'_{\mathfrak{R}}$ und einer beliebigen Umgebung U' von x' in E' existiert nämlich eine stetige Funktion $f' | E'_{\mathfrak{R}}$ mit $f'(x') = 1$ und $f'(y') = 0$ für alle $y' \in U'$. Nach Lemma 4 und (31) existiert zu f' eine Funktion $g' \in \mathfrak{R}'$ mit $|f'(z') - g'(z')| \leq 1/3$ für alle $z' \in E'_{\mathfrak{R}}$. g' hat die Gestalt $g' = \sum \alpha_\nu \chi'_{\bar{A}_\nu} (\alpha_\nu \in \mathsf{P}, \ A_\nu \in \mathfrak{B})$. Wegen $f'(x') = 1$ ist $g'(x') \neq 0$; folglich liegt x' in einer Menge $\bar{A}_{\nu_0}$ und es ist $|1 - \alpha_{\nu_0}| \leq 1/3$. Da dann nicht $|\alpha_{\nu_0}| \leq 1/3$ sein kann, gilt $\bar{A}_{\nu_0} \subseteq U'$. Also ist $\overline{\mathfrak{B}}$ eine offene Basis von $E'_{\mathfrak{R}}$. Ordnet man jedem Punkt $x' \in E'_{\mathfrak{R}}$ als Bild $\psi(x')$ in $U_{\mathfrak{B}}$ den aus allen Mengen $A \in \mathfrak{B}$ mit $x' \in \bar{A}$ bestehenden Ultrafilter in $\mathfrak{B}$ zu, so ist

[31] Vgl. PAUC [*18*], S. 33.

$\psi(\bar{A}) = \Phi A$ für jede Menge $A \in \mathfrak{B}$. Hieraus und den bewiesenen Eigenschaften von $E'_{\mathfrak{R}}$ folgt, daß ψ die behauptete Homöomorphie von $E'_{\mathfrak{R}}$ auf $U_{\mathfrak{B}}$ ist, welche E punktweise festläßt.

Wir führen dieses Beispiel weiter und geben eine *Kennzeichnung des σ-additiven und rein-endlich-additiven Teils einer additiven Mengenfunktion* an. Auf der Boóleschen Mengenalgebra $\mathfrak{B}$ sei nämlich eine nicht negative, (endlich-)additive Funktion φ definiert. Wir übertragen sie von $\mathfrak{B}$ auf die zu $\mathfrak{B}$ isomorphe Boolesche Algebra $\overline{\mathfrak{B}}$ der abgeschlossenen Hüllen $\bar{A}$ von Mengen $A \in \mathfrak{B}$ in $E'_{\mathfrak{R}}$, indem wir $\varphi'(\bar{A}) = \varphi(A)$ für jedes $A \in \mathfrak{B}$ bzw. $\bar{A} \in \overline{\mathfrak{B}}$ setzen. Dann ist $\varphi' | \overline{\mathfrak{B}}$ trivialerweise σ-additiv; aus φ' leitet sich daher das auf dem System aller Teilmengen von $E'_{\mathfrak{R}}$ definierte äußere Maß $\overline{\varphi'}$ und innere Maß $\underline{\varphi'}$ ab. Wir behaupten:

Bezeichnet φ_c den σ-additiven und φ_p den rein-endlich-additiven Teil von $\varphi | \mathfrak{B}$, so gilt

$$\varphi_c(A) = \overline{\varphi'}(A) \quad \textit{und} \quad \varphi_p(A) = \underline{\varphi'}(\bar{A} - A)$$

für jede Menge $A \in \mathfrak{B}$[32]).

Beweis. Der additiven Funktion $\varphi | \mathfrak{B}$ entspricht in umkehrbar eindeutiger Weise ein positives, lineares Funktional $I | \mathfrak{R}$: Für eine Funktion $g = \sum \alpha_\nu \chi_{A_\nu}$ aus $\mathfrak{R}$ ist $I(g) = \sum \alpha_\nu \varphi(A_\nu)$. Aus $I | \mathfrak{R}$ leitet sich das Funktional $I' | \mathfrak{R}'$ und aus diesem der an $E'_{\mathfrak{R}}$ adaptierte Inhalt $\mu' | \mathfrak{Q}'$ ab. Aus der Definition von $\mu' | \mathfrak{Q}'$ folgt, da jede Funktion $g' \in \mathfrak{R}'$ Linearkombination von charakteristischen Funktionen $\chi'_{\bar{A}}$ $(\bar{A} \in \overline{\mathfrak{B}})$ ist, daß $\mu' | \mathfrak{Q}'$ die kleinste Erweiterung von $\varphi' | \overline{\mathfrak{B}}$ zu einem vollständigen Inhalt ist. Dann ist aber $\overline{\varphi'} = \overline{\mu'}$ sowie $\underline{\varphi'} = \underline{\mu'}$. Die Behauptung folgt daher aus Satz 18.

Aus den bewiesenen Formeln für φ_c und φ_p folgen nun notwendige und hinreichende Bedingungen dafür, daß φ σ-additiv bzw. rein-endlich-additiv ist; diese Bedingungen sind zu denen im Korollar zu Satz 18 formulierten analog. Insbesondere erweist sich φ dann und nur dann als σ-additiv, wenn $\underline{\varphi'}(\bar{E} - E) = 0$ ist. Dieses Kriterium findet sich bereits bei PAUC [18], S. 35.

Literatur

[1] ARENS, R.: Representation of functionals by integrals. Duke math. J. **17**, 499—506 (1950). — [2] AUMANN, G.: Reelle Funktionen. Berlin-Göttingen-Heidelberg: Springer 1954. — [3] BAUER, H.: Eine Rieszsche Bandzerlegung im Raum der Bewertungen eines Verbandes. Sitzgsber. math.-nat. Kl. Bayer. Akad. Wiss. München **1953**, 89—117 (1954). — [4] BAUER, H.: Caractérisation topologique de la partie complètement additive et de la partie purement additive d'une fonction additive d'ensemble. C. r. Acad. Sci. Paris **238**, 1771—1773 (1954). — [5] BAUER, H.: Reguläre und singuläre Abbildungen eines distributiven Verbandes in einen vollständigen Vektorverband, welche der Funktionalgleichung $f(x \vee y) + f(x \wedge y) = f(x) + f(y)$ genügen. J. reine angew. Math. **194**, 141—179 (1955). — [6] BAUER, H.: Darstellung additiver Funktionen auf Booleschen Algebren als Mengenfunktionen. Arch. der Math. **6**, 215—222 (1955). — [7] BAUER, H.: Zur Theorie des Riemann-Integrals in lokal kompakten Räumen. Sitzgsber. math.-nat. Kl. Bayer. Akad. Wiss. München **1955**, 187—208 (1956). — [8] BOURBAKI, N.:

[32]) Da $E'_{\mathfrak{R}} = U_{\mathfrak{B}}$ ist, erweist sich jetzt dieses in [4] unter anderem von uns angekündigte Resultat als Spezialfall von Satz 18.

Intégration, Chap. I—IV. Actual. Scient. et Ind. **1175**, Paris 1952. — [*9*] GELFAND, I., u. G. ŠILOV: Über verschiedene Methoden der Einführung der Topologie in die Menge der maximalen Ideale eines normierten Ringes. Rec. math. Moscou, n. ser. **9**, 25—38 (1941). — [*10*] HAUPT, O., et CHR. PAUC: Mesure et topologie adaptées. Espaces mesurés topologiques. C. r. Acad. Sci. Paris **230**, 711—712 (1950). — [*11*] HAUPT, O., u. CHR. PAUC: Bemerkungen über Inhalte und Maße in lokal bikompakten Räumen. Akad. Wiss. Lit. Mainz, Abh. math.-nat. Kl. **1955**, 189—218 (1956). — [*12a*] HAUPT-AUMANN-PAUC: Differential- und Integralrechnung, Bd. I, 2. Aufl. Göschens Lehrbücherei **24**, Berlin 1948. — [*12b*] HAUPT-AUMANN-PAUC: Differential- und Integralrechnung, Bd. III, 2. Aufl. Göschens Lehrbücherei **26**, Berlin 1955. — [*13*] HEWITT, E.: Integral representation of certain linear functionals. Arkiv för Matematik **2**, 269—282 (1951). — [*14*] LOOMIS, L. H.: Linear functionals and content. Amer. J. Math. **76**, 168—182 (1954). — [*15*] NAKANO, H.: On the product of relative spectra. Ann. of Math., II. s. **49**, 281—315 (1948). — [*16*] NÖBELING, G., u. H. BAUER: Über die Erweiterungen topologischer Räume. Math. Ann. **130**, 20—45 (1955). — [*17*] NÖBELING, G., u. H. BAUER: Allgemeine Approximationskriterien mit Anwendungen. J.-Ber. Deutsch. Math.-Verein. **58**, 54—72 (1955). — [*18*] PAUC, CHR.: Darstellungs- und Struktursätze für Boolesche Verbände und σ-Verbände. Arch. der Math. **1**, 29—41 (1948). — [*19*] PAUC, CHR.: Intégrale de partition et intégrale topologique. Familles dérivantes topologiques. C. r. Acad. Sci. Paris **230**, 810—811 (1950). — [*20*] STONE, M. H.: Applications of the theory of Boolean rings to general topology. Trans. Amer. math. Soc. **41**, 375—481 (1937). — [*21*] STONE, M. H.: Notes on integration, I—IV. Proc. Nat. Acad. Sci. U.S.A. **34**, 336—342, 447—455, 483—490 (1948); **35**, 50—58 (1949). — [*22*] WOODBURY, M. A.: A decomposition theorem for finitely additive set functions (Preliminary report). Bull. Amer. math. Soc. **56**, 171 (1950). — [*23*] YOSIDA, K., and E. HEWITT: Finitely additive measures. Trans. Amer. math. Soc. **72**, 46—66 (1952).

Erlangen, Mathematisches Institut der Universität

(Eingegangen am 2. März 1956)

SONNER, H.
Math. Zeitschr. Bd. 65, S. 483—493 (1956)

Über die Zurückführung partieller Differentialgleichungen auf gewöhnliche

Von
HANS SONNER

Einleitung. Die Theorie der gewöhnlichen Differentialgleichungen in *vollständigen normierten Räumen*, wie sie etwa bei BOURBAKI [4] durchgeführt wird, ist unter passenden Umständen geeignet, die Lösung partieller Differentialgleichungen auf die von gewöhnlichen zurückzuführen. Der Gedanke besteht darin, erstens jeder partiellen Differentialgleichung von $p+1$ unabhängigen reellen Veränderlichen eine gewöhnliche Differentialgleichung zuzuordnen, deren Lösungen Funktionen mit *Werten im Raume der Funktionen* von p reellen Variablen sind, und zweitens den Raum dieser Funktionen durch Einführung einer Norm zu einem vollständigen normierten Raum zu machen.

Nun lehren einfachste Beispiele (vgl. § 1, Bemerkung), daß mit jeder Funktion dieses Raumes auch deren sämtliche partiellen Ableitungen zum Raume gehören. Bei der Bildung der Norm muß dies berücksichtigt werden, was dazu führt, von vornherein unendlich oft differenzierbare, ableitungsbeschränkte Funktionen zu betrachten, d.h. Funktionen, deren sämtliche Ableitungen eine gemeinsame Schranke besitzen.

Wegen der Normiertheit unseres Funktionenraumes können wir Abbildungen von Intervallen der reellen Zahlengeraden in diesen Raum betrachten und von Ableitungen sprechen. In diesem Sinn wird in § 1 jeder Differentialgleichung, die wir uns in Normalform gegeben denken, eine gewöhnliche Differentialgleichung zugeordnet; ferner wird der Begriff der Lösung einer partiellen Differentialgleichung präzisiert. Das Studium des Raumes der unendlich oft differenzierbaren, ableitungsbeschränkten Funktionen ist dem § 2 vorbehalten. In § 3 wird gezeigt, daß sich die „starken" Lösungen der partiellen Gleichung aus den Lösungen der zugehörigen gewöhnlichen Gleichung gewinnen lassen. Um zu Existenz- und Eindeutigkeitsaussagen zu gelangen, wird in § 4 die LIPSCHITZ-Bedingung in verschärfter Form eingeführt. Mit ihrer Hilfe wird in § 5 ein Existenzsatz formuliert, der dem entsprechenden Satz bei gewöhnlichen Gleichungen wörtlich nachgebildet ist. Nach demselben Muster würden sich (was hier nicht durchgeführt ist) auch Sätze über die Stetigkeit bezüglich eines Parameters und über die Abhängigkeit von den Anfangsbedingungen übertragen lassen. An zwei Beispielen wird das Verfahren erläutert. Der Schluß bringt eine Abgrenzung des in der vorliegenden Arbeit dargelegten Verfahrens gegenüber dem von CAUCHY und KOWALEWSKI herrührenden Verfahren.

Zeichenerklärung. Im folgenden bezeichne $\mathbf{N}$ (bzw. $\mathbf{N}^*$) die Menge der natürlichen Zahlen (bzw. $\neq 0$), $\mathbf{R}$ die Menge der reellen Zahlen, $\mathbf{C}$ die Menge

der komplexen Zahlen, $A \times B$ bzw. $\prod_{0 \le i \le p} I_i$ kartesische Produkte, $I - A$ das Komplement von A bezüglich I, D_i die partielle Ableitung nach der i-ten Variablen, etwa $\langle a, b \rangle$ das links abgeschlossene, rechts offene Intervall mit den Endpunkten a und b. Fette Buchstaben bedeuten Vektoren. Ein Mengentripel (G, A, B) mit $G \subset A \times B$ nennen wir (nach BOURBAKI) Korrespondenz, eine Korrespondenz (G, A, B), bei der es zu jedem $x \in A$ ein und nur ein $y \in B$ mit $(x, y) \in G$ gibt, Funktion. Setzen wir $f = (G, A, B,)$ so wird dieses y mit $f(x)$ bezeichnet. Ist der Graph G einer Funktion (G, A, B) Graph einer Relation der Gestalt „$x \in A$ und $y \in B$ und $y = T$", wo T ein Term ist, in dem x vorkommt, so schreiben wir statt (G, A, B) eindrucksvoller $x \to T$ $(x \in A,\ T \in B)$ oder im Zusammenhang auch kurz $x \to T$. Ist u eine von $p+1$ Variablen abhängige Funktion, so bezeichne u_{t_0} die der Identität $u_{t_0}(t_1, \ldots, t_p) = u(t_0, t_1, \ldots, t_p)$ genügende partielle Funktion. $f \circ g$ sei die der Identität $(f \circ g)(x) = f(g(x))$ genügende Komposition der Funktionen f und g.

§ 1. Der Begriff der Lösung

Gegeben seien eine Familie $(I_i)_{0 \le i \le p}$ von $p+1$ nicht auf einen Punkt reduzierten, in $\mathbf{R}$ enthaltenen Intervallen, ein vollständiger normierter Raum E über $\mathbf{R}$ und eine Familie $(H_i)_{0 \le i \le p}$ von $p+1$ offenen, nicht leeren Teilen von E. Sei ferner $\mathfrak{f}$ eine Abbildung von $\prod_{0 \le i \le p} I_i \times \prod_{0 \le i \le p} H_i$ in E. Wir sagen, daß eine in $\prod_{0 \le i \le p} I_i$ erklärte Funktion $\mathfrak{u}$ mit Werten in E *Lösung* der Differentialgleichung

$$D_0 \mathfrak{x} = \mathfrak{f}(t_0, t_1, \ldots, t_p, \mathfrak{x}, D_1 \mathfrak{x}, \ldots, D_p \mathfrak{x}) \tag{1}$$

ist, wenn erstens $\mathfrak{u}$ stetig ist und stetige partielle Ableitungen $D_1 \mathfrak{u}, \ldots, D_p \mathfrak{u}$ besitzt, und zweitens ein abzählbarer Teil A von I_0 existiert, so daß $\mathfrak{u}$ in jedem Punkt t des Produktes $\prod_{0 \le i \le p} I_i'$ mit $I_0' = I_0 - A$ und $I_j' = I_j$ $(1 \le j \le p)$ eine partielle Ableitung $D_0 \mathfrak{u}(t)$ mit $D_0 \mathfrak{u}(t) = \mathfrak{f}\big(t, \mathfrak{u}(t), D_1 \mathfrak{u}(t), \ldots, D_p \mathfrak{u}(t)\big)$ hat. Falls für A der *leere* Teil von I_0 gewählt werden kann, sprechen wir von einer *strengen* Lösung der Gl. (1) in $\prod_{0 \le i \le p} I_i$. (1) heißt genauer „partielle Differentialgleichung erster Ordnung in Normalform". Man beachte, daß die hier gegebene Definition der Lösung im Falle $p = 0$ mit der üblichen für gewöhnliche Differentialgleichungen übereinstimmt.

Die der Variablen t_0 eingeräumte Sonderstellung erlaubt in vielen Fällen, die Lösung einer partiellen Differentialgleichung von $p+1$ Veränderlichen auf die einer gewöhnlichen Differentialgleichung zurückzuführen. Bezeichne zu diesem Zweck $\tilde{E}$ die Menge aller in $Q = \prod_{1 \le i \le p} I_i$ erklärten unendlich oft differenzierbaren, ableitungsbeschränkten Funktionen $\tilde{\mathfrak{x}}$ mit Werten in E, $\|\tilde{\mathfrak{x}}\|$ das Supremum der Menge aller Zahlen $\|\tilde{\mathfrak{x}}(\tau)\|$ $(\tau \in Q)$, $[\tilde{\mathfrak{x}}]$ das (endliche) Supremum der Menge aller Zahlen $\|D_1^{n_1} \ldots D_p^{n_p} \tilde{\mathfrak{x}}\|$ $\big((n_1, \ldots, n_p) \in \mathbf{N}^p\big)$ und schließlich $\tilde{H}$ die Menge aller Funktionen $\tilde{\mathfrak{x}} \in \tilde{E}$, für die die abgeschlossene Hülle $\overline{Q'}$ des

Bildes Q' durch $\tau \to \left(\tilde{\mathfrak{x}}(\tau), D_1\tilde{\mathfrak{x}}(\tau), \ldots, D_p\tilde{\mathfrak{x}}(\tau)\right)$ von Q kompakter Teil des Produktraumes $\prod_{0\leq i\leq p} H_i$ ist. Zu $t_0 \in I_0$ und $\tilde{\mathfrak{x}} \in \tilde{H}$ bilden wir die Funktion

$$\tau \to \mathfrak{f}\left(t_0, \tau, \tilde{\mathfrak{x}}(\tau), D_1\tilde{\mathfrak{x}}(\tau), \ldots, D_p\tilde{\mathfrak{x}}(\tau)\right) \quad \left(\tau \in Q, \mathfrak{f}(t_0, \tau, \tilde{\mathfrak{x}}(\tau), D_1\tilde{\mathfrak{x}}(\tau), \ldots, D_p\tilde{\mathfrak{x}}(\tau)) \in E\right)$$

und geben ihr den Namen $\tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})$. Das Zeichen $\tilde{\mathfrak{f}}$ behalten wir der Korrespondenz $(\Phi, I_0 \times \tilde{H}, \tilde{E})$ vor, wobei Φ die Menge aller Paare $\left((t_0, \tilde{\mathfrak{x}}), \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})\right)$ mit $(t_0, \tilde{\mathfrak{x}}) \in I_0 \times \tilde{H}$ und $\tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}}) \in \tilde{E}$ ist. Falls $\tilde{\mathfrak{x}} \to [\tilde{\mathfrak{x}}]$ die Menge $\tilde{E}$ zu einem vollständigen normierten Raum macht, können wir Abbildungen eines Intervalls der reellen Zahlengeraden in diesen Raum betrachten und Ableitungen erklären. In diesem Sinn betrachten wir statt (1) die (gewöhnliche) Differentialgleichung

$$\tilde{\mathfrak{x}}' = \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}}). \tag{2}$$

Für eine Lösung $\mathfrak{u}$ von (1) bezeichne $\tilde{\mathfrak{u}}$ die Korrespondenz $(\Psi, I_0, \tilde{E})$, wobei Ψ die Menge aller Paare $(t_0, \mathfrak{u}_{t_0})$ mit $t_0 \in I_0$ und $\mathfrak{u}_{t_0} \in \tilde{E}$ ist. Diese kann man, wie wir sehen werden, unter geeigneten Umständen als Lösung von (2) deuten. Wir sagen, eine Lösung $\mathfrak{u}$ von (1) in $\prod_{0\leq i\leq p} I_i$ gehe in die Funktion $\tilde{\mathfrak{x}}_* \in \tilde{H}$ im Punkt $t_* \in I_0$ über, wenn $\mathfrak{u}_{t_*} = \tilde{\mathfrak{x}}_*$ ist.

Beispiele. 1. Die homogene lineare Differentialgleichung

$$D_0 x = a(t_0, t_1) \cdot x + b(t_0, t_1) \cdot D_1 x$$

besitzt im Falle konstanter Koeffizienten die Lösung $(t_0, t_1) \to e^{at_0} \cdot \tilde{x}_*(t_1 + bt_0)$, die wir u nennen. Diese geht in die Funktion $\tilde{x}_*$ im Punkt 0 über. Die zugeordnete Gleichung $\tilde{x}' = a_{t_0} \cdot \tilde{x} + b_{t_0} \cdot D_1\tilde{x}$ besitzt im Falle konstanter Koeffizienten die zugeordnete Lösung $t_0 \to u_{t_0}$. Diese nimmt den Wert (lies Funktion) $\tilde{x}_*$ im Punkt 0 an.

2. Die hyperbolische Gleichung $\frac{\partial^2 z}{\partial x^2} - \frac{\partial^2 z}{\partial y^2} = a(x, y)$ läßt sich mit unserer Bezeichnungsweise als lineares System erster Ordnung in der Normalform

$$D_0\xi_1 = \xi_2, \quad D_0\xi_2 = D_1\xi_3 + a(t_0, t_1), \quad D_0\xi_3 = D_1\xi_2$$

schreiben. Die Funktion $\mathfrak{f}$ lautet $(t_0, t_1, (\xi_{0\varrho})_{1\leq\varrho\leq 3}, (\xi_{1\varrho})_{1\leq\varrho\leq 3}) \to (\xi_{02}, \xi_{13} + a(t_0, t_1), \xi_{12})$. Dem partiellen Differentialgleichungssystem ist das gewöhnliche Differentialgleichungssystem

$$\tilde{\xi}_1' = \tilde{\xi}_2, \quad \tilde{\xi}_2' = D_1\tilde{\xi}_3 + a_{t_0}, \quad \tilde{\xi}_3' = D_1\xi_2$$

zugeordnet. $\tilde{\mathfrak{f}}$ hat das Aussehen $(t_0, (\tilde{\xi}_\varrho)_{1\leq\varrho\leq 3}) \to (\tilde{\xi}_2, D_1\tilde{\xi}_3 + a_{t_0}, D_1\tilde{\xi}_2)$. Dabei haben wir $\mathfrak{x} = (\xi_\varrho)_{1\leq\varrho\leq 3}$, $\mathfrak{x}_0 = (\xi_{0\varrho})_{1\leq\varrho\leq 3}$, $\mathfrak{x}_1 = (\xi_{1\varrho})_{1\leq\varrho\leq 3}$, $\tilde{\mathfrak{x}} = (\tilde{\xi}_\varrho)_{1\leq\varrho\leq 3}$ gesetzt.

Bemerkung. Zu den partiellen Differentialgleichungen $D_0\mathfrak{x} = D_j\mathfrak{x}$ $(1 \leq j \leq p)$ gehören die gewöhnlichen Differentialgleichungen $\tilde{\mathfrak{x}}' = D_j\tilde{\mathfrak{x}}$ $(1 \leq j \leq p)$. Also ist $\tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}}) = D_j\tilde{\mathfrak{x}}$ $(1 \leq j \leq p)$. Soll mit $\tilde{\mathfrak{x}}$ auch $\tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})$ Element von $\tilde{E}$ sein, so muß notwendig mit $\tilde{\mathfrak{x}}$ auch die partielle Ableitung $D_j\tilde{\mathfrak{x}}$ $(1 \leq j \leq p)$ zu $\tilde{E}$ gehören; d.h. die Funktionen von $\tilde{E}$ sind unendlich oft differenzierbar.

§ 2. Der Raum der unendlich oft differenzierbaren, ableitungsbeschränkten Funktionen

Satz 1. *Sei E ein vollständiger normierter Raum über* R. *Dann ist der mit der Norm* $[\tilde{\mathfrak{x}}]$ *versehene Vektorraum* $\tilde{E}$ *ein vollständiger normierter, nur aus Funktionen mit stetigen partiellen Ableitungen bestehender Raum über* R.

Die Vektor- und Normeigenschaften ergeben sich aus den entsprechenden Eigenschaften von E und der Linearität der Ableitungen. Es genügt also, die Vollständigkeit nachzuweisen. Sei zu diesem Zweck $\mathfrak{F}$ ein CAUCHY-Filter (bezüglich der Norm $[\tilde{\mathfrak{x}}]$) auf $\tilde{E}$. Offenbar ist dann auch für jedes p-Tupel $(n_1, \ldots, n_p)$ von N^p das System $\mathfrak{B}^{(n_1, \ldots, n_p)}$, das aus $\mathfrak{F}$ entsteht, indem man in jeder Menge von $\mathfrak{F}$ die Funktionen $\tilde{\mathfrak{x}}$ durch die Ableitungen $D_p^{n_p} \ldots D_1^{n_1} \tilde{\mathfrak{x}}$ ersetzt, Basis eines CAUCHY-Filters $\mathfrak{F}^{(n_1, \ldots, n_p)}$ auf der Menge $\mathfrak{B}(Q, E)$ aller beschränkten Abbildungen von Q in E (bezüglich der Norm $\|\tilde{\mathfrak{x}}\|$). Wegen der Vollständigkeit von E konvergieren diese Filter $\mathfrak{F}^{(n_1, \ldots, n_p)}$ gleichmäßig gegen gewisse (beschränkte) Funktionen $\tilde{\mathfrak{x}}^{(n_1, \ldots, n_p)}$, die notwendig mit den Ableitungen $D_1^{n_1} \ldots D_p^{n_p} \tilde{\mathfrak{x}}_0$ von $\tilde{\mathfrak{x}}_0 = \tilde{\mathfrak{x}}^{(0, \ldots, 0)}$ übereinstimmen.

Um die Endlichkeit von $[\tilde{\mathfrak{x}}_0]$ und die Konvergenz von $\mathfrak{F}$ gegen $\tilde{\mathfrak{x}}_0$ bezüglich der Norm $[\tilde{\mathfrak{x}}]$ zu zeigen, wähle man zu $\varepsilon > 0$ ein $A \in \mathfrak{F}$ so, daß für alle Paare $\tilde{\mathfrak{x}}, \tilde{\mathfrak{y}}$ von A und alle p-Tupel $(n_1, \ldots, n_p)$ von N^p die Ungleichung $\|D_1^{n_1} \ldots D_p^{n_p} (\tilde{\mathfrak{x}} - \tilde{\mathfrak{y}})\| \leq \varepsilon$ gilt. Schafft man nach Anwendung der Dreiecksungleichung das $\tilde{\mathfrak{y}}$ enthaltende Glied auf die rechte Seite und beachtet die gleichmäßige Konvergenz der Filter $\mathfrak{F}^{(n_1, \ldots, n_p)}$, so kommt für ein festes $\tilde{\mathfrak{y}}$ schließlich $[\tilde{\mathfrak{x}}_0] \leq \varepsilon + [\tilde{\mathfrak{y}}]$, also die Endlichkeit von $\tilde{\mathfrak{x}}_0$. Geht man dagegen in der obigen Ungleichung sofort zum Limes über, so gilt für alle $\tilde{\mathfrak{y}} \in A$ die Relation $[\tilde{\mathfrak{x}}_0 - \tilde{\mathfrak{y}}] \leq \varepsilon$, d.h. die behauptete Konvergenz.

Die Stetigkeit der partiellen Ableitung $\tilde{\mathfrak{y}} = D_1^{n_1} \ldots D_p^{n_p} \tilde{\mathfrak{x}}$ ergibt sich aus der Abschätzung

$$\|\tilde{\mathfrak{y}}(\tau') - \tilde{\mathfrak{y}}(\tau)\| = \Big\| \sum_{1 \leq k \leq p} \big(\tilde{\mathfrak{y}}(t_1', \ldots, t_{k-1}', t_k', t_{k+1}, \ldots, t_p) - \tilde{\mathfrak{y}}(t_1', \ldots, t_{k-1}', t_k, t_{k+1}, \ldots, t_p)\big) \Big\|$$

$$\leq \sum_{1 \leq k \leq p} |t_k' - t_p| \cdot \big(\|D_k \tilde{\mathfrak{y}}\| + o_k(t' - t)\big) \leq \{[\tilde{\mathfrak{x}}] + 1\} \sum_{1 \leq k \leq p} |t_k' - t_k|,$$

falls τ' hinreichend nahe bei τ liegt.

Satz 2. (a) *Für jede stetige lineare Abbildung* U *von E in E ist mit* $\tilde{\mathfrak{x}}$ *auch* $\mathsf{U} \circ \tilde{\mathfrak{x}}$ *eine Funktion von* $\tilde{E}$, *und es gilt* $[\mathsf{U} \circ \tilde{\mathfrak{x}}] \leq \|\mathsf{U}\| \cdot [\tilde{\mathfrak{x}}]$.

(b) *Für jedes p-Tupel $(n_1, \ldots, n_p)$ von* N^p *ist mit* $\tilde{\mathfrak{x}}$ *auch* $D_1^{n_1} \ldots D_p^{n_p} \tilde{\mathfrak{x}}$ *eine Funktion von* $\tilde{E}$, *und es gilt* $[D_1^{n_1} \ldots D_p^{n_p} \tilde{\mathfrak{x}}] \leq [\tilde{\mathfrak{x}}]$.

(a) Mit $\tilde{\mathfrak{x}}$ ist wegen der Stetigkeit der linearen Abbildung auch $\mathsf{U} \circ \tilde{\mathfrak{x}}$ unendlich oft differenzierbar und es gilt $D_1^{n_1} \ldots D_p^{n_p} (\mathsf{U} \circ \tilde{\mathfrak{x}}) = \mathsf{U} \circ (D_1^{n_1} \ldots D_p^{n_p} \tilde{\mathfrak{x}})$, also wegen der Beschränktheit einer stetigen linearen Abbildung

$$\|D_1^{n_1} \ldots D_p^{n_p} (\mathsf{U} \circ \tilde{\mathfrak{x}})\| \leq \|\mathsf{U}\| \cdot \|D_1^{n_1} \ldots D_p^{n_p} \tilde{\mathfrak{x}}\| \quad \text{oder} \quad [\mathsf{U} \circ \tilde{\mathfrak{x}}] \leq \|\mathsf{U}\| \cdot [\tilde{\mathfrak{x}}].$$

(b) Ergibt sich sofort aus der Definition von $[\tilde{\mathfrak{x}}]$.

Satz 3. *Sei* $(H_i)_{0 \leq i \leq p}$ *eine Familie von offenen Teilen von E. Dann ist* $\tilde{H}$ *ein offener Teil von* $\tilde{E}$.

Wir wählen zu vorgegebenem $\tilde{\mathfrak{x}} \in \tilde{H}$ ein $\varrho > 0$ so, daß die abgeschlossenen Kugeln mit dem Radius ϱ und dem Zentrum $(\tilde{\mathfrak{x}}(\tau), D_1\tilde{\mathfrak{x}}(\tau), \ldots, D_p\tilde{\mathfrak{x}}(\tau))$ $(\tau \in Q)$ noch in $\prod_{0 \le i \le p} H_i$ liegen, was nach Definition von $\tilde{H}$ und der Offenheit von $\prod_{0 \le i \le p} H_i$ möglich ist. Dann gehört, wie man leicht nachrechnet, die aus allen Funktionen $\tilde{\mathfrak{y}} \in \tilde{E}$ mit $[\tilde{\mathfrak{x}} - \tilde{\mathfrak{y}}] \le \varrho$ bestehende Umgebung von $\tilde{\mathfrak{x}}$ zu $\tilde{H}$.

Bemerkung. Wie das Beispiel $E = \mathsf{R}$, $I_i = \langle 0, 1 \rangle$ $(0 \le i \le p)$, $H_0 = \rangle 0, 1 \langle$, $H_1 = \rangle 2, 3 \langle$, $H_j = \mathsf{R}\,(2 \le j \le p)$ lehrt, gibt es Familien $(H_i)_{0 \le i \le p}$ von nichtleeren Mengen H_i, die zu leerem $\tilde{H}$ führen.

§ 3. Zusammenhang der Lösungen

Satz 4. *Sei t_* ein Punkt von I_0, $\tilde{\mathfrak{x}}_*$ eine Funktion von $\tilde{H}$, $\mathfrak{u}$ eine in $\prod_{0 \le i \le p} I_i$ erklärte Funktion mit Werten in E. Ist die zugehörige Korrespondenz $\tilde{\mathfrak{u}}$ Lösung der Gl. (2) in I_0 und nimmt sie den Wert $\tilde{\mathfrak{x}}_*$ im Punkt t_* an, so ist $\mathfrak{u}$ Lösung der Gl. (1) in $\prod_{0 \le i \le p} I_i$ und geht in die Funktion $\tilde{\mathfrak{x}}_*$ im Punkt t_* über.*

(a) Die Stetigkeit von $\mathfrak{u}$ ergibt sich aus der Abschätzung

$$\|\mathfrak{u}(t_0', \tau) - \mathfrak{u}(t_0, \tau)\| \le [\mathfrak{u}_{t_0'} - \mathfrak{u}_{t_0}] + \|\mathfrak{u}(t_0, \tau') - \mathfrak{u}(t_0, \tau)\|;$$

das erste Glied der rechten Seite kann wegen der Stetigkeit der Funktion $t_0 \to \mathfrak{u}_{t_0}$ im Punkt t_0, das zweite wegen der Stetigkeit der Funktion $\mathfrak{u}_{t_0}$ im Punkt τ beliebig klein gemacht werden.

(b) Da $t_0 \to \mathfrak{u}_{t_0}$ als Lösung von (2) das Intervall I_0 in den Raum $\tilde{E}$ der unendlich oft differenzierbaren, ableitungsbeschränkten Funktionen abbildet, ist $\mathfrak{u}_{t_0}$ für alle $t_0 \in I_0$ unendlich oft differenzierbar und ableitungsbeschränkt.

(c) Bei festem $\tau \in Q$ ist $t_0 \to \mathfrak{u}(t_0, \tau)$ in jedem Punkt t_0 von I_0 differenzierbar, in dem es auch $t_0 \to \mathfrak{u}_{t_0}$ ist; denn zu vorgegebenem $\varepsilon > 0$ gibt es eine Umgebung V von t_0, so daß für alle t_0' von V die Relation

$$\left[\frac{\tilde{\mathfrak{u}}(t_0') - \tilde{\mathfrak{u}}(t_0)}{t_0' - t_0} - \tilde{\mathfrak{u}}'(t_0)\right] \le \varepsilon$$

gilt. Daraus folgt für alle t_0' von V die Ungleichung

$$\left\| \frac{\mathfrak{u}(t_0', \tau) - \mathfrak{u}(t_0, \tau)}{t_0' - t_0} - D_0\mathfrak{u}(t_0, \tau) \right\| \le \varepsilon,$$

wobei $D_0\mathfrak{u}(t_0, \tau)$ für den Wert der Funktion $\tilde{\mathfrak{u}}'(t_0)$ im Punkt τ geschrieben ist. Nun gilt in jedem Punkt t_0 des Komplements bezüglich I_0 einer abzählbaren Menge A die Gleichung $\tilde{\mathfrak{u}}'(t_0) = \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{u}}(t_0))$, also in jedem Punkt t des Produktes $\prod_{0 \le i \le p} I_i'$ mit $I_0' = I_0 - A$ und $I_i' = I_j\,(1 \le j \le p)$ nach dem eben Bewiesenen $D_0\mathfrak{u}(t) = \mathfrak{f}(t, \mathfrak{u}(t), D_1\mathfrak{u}(t), \ldots, D_p\mathfrak{u}(t))$.

(d) Es ist $\mathfrak{u}_{t_*}$ gleich dem Wert der Lösung $t_0 \to \mathfrak{u}_{t_0}$ für den Punkt t_*, also gleich $\mathfrak{x}_*$.

Bemerkung. Daß nicht jede Lösung $\mathfrak{u}$ von (1) zu einer Lösung $\tilde{\mathfrak{u}}$ von (2) führt, lehrt die Differentialgleichung $D_0 x = t_1 x$. Wir setzen, um etwas

Bestimmtes vor Augen zu haben, $I_0 = I_1 = \langle 1, 2\rangle$, $E = H_0 = H_1 = \mathbf{C}$. Zu ihr gehört die gewöhnliche Gleichung $\tilde{x}' = \tilde{f}(t_0, \tilde{x}) = \varphi \cdot \tilde{x}$, wobei φ die kanonische Abbildung von I_1 in $\mathbf{C}$ bedeutet. Offensichtlich ist $(t_0, t_1) \to e^{t_0 t_1}$ strenge Lösung der partiellen Gleichung und geht in die Funktion $t_1 \to e^{t_1}$ im Punkt $t_* = 1$ über; wir geben ihr den Namen u. Dagegen ist $\tilde{u}$ nicht Lösung der gewöhnlichen Gleichung; denn es ist $\|D_1^n u_{t_0}\| = t_0^n \cdot e^{2t_0}$, also $[u_{t_0}] = e^2$, falls $t_0 = 1$ ist, $= \infty$, falls $t_0 > 1$ ist. Es gehört also das Paar (t_0, u_{t_0}) nur für $t_0 = t_* = 1$ zum Graph von $\tilde{u}$.

Definition 1. *Wir nennen* starke *Lösung von* (1) *jede Lösung* $\mathfrak{u}$ *von* (1), *für die die zugehörige Korrespondenz* $\tilde{\mathfrak{u}}$ *Lösung der Gl.* (2) *ist.*

Satz 5. *Sei* $\mathfrak{u}$ *eine starke Lösung von* (1). *Dann ist für jedes feste* $t_0 \in I$ *die Funktion* $\mathfrak{u}_{t_0}$ *unendlich oft differenzierbar und ableitungsbeschränkt. Ferner ist die Familie der Funktionen* $t_0 \to D_1^{n_1} \dots D_p^{n_p} \mathfrak{u}(t_0, \tau)$ $(\tau \in Q, (n_1, \dots, n_p) \in \mathbf{N}^p)$ *gleichgradig stetig.*

Die erste Behauptung ergibt sich unmittelbar aus Definition 1, die zweite auf Grund der Abschätzung

$$\|D_1^{n_1} \dots D_p^{n_p} \mathfrak{u}(t_0', \tau) - D_1^{n_1} \dots D_p^{n_p} \mathfrak{u}(t_0, \tau)\| \leq \|D_1^{n_1} \dots D_p^{n_p} \mathfrak{u}_{t_0'} - D_1^{n_1} \dots D_p^{n_p} \mathfrak{u}_{t_0}\| \leq$$
$$\leq [\mathfrak{u}_{t_0'} - \mathfrak{u}_{t_0}]$$

und der Stetigkeit von $t_0 \to \mathfrak{u}_{t_0}$.

§ 4. Die verschärfte LIPSCHITZ-Bedingung

Um zu Existenz- und Eindeutigkeitsaussagen zu gelangen, ist es nützlich, der Funktion $\mathfrak{f}$ gewisse Bedingungen aufzuerlegen. Diese drücken sich am einfachsten in der Korrespondenz $\tilde{\mathfrak{f}}$ aus.

Definition 2. *Man sagt, daß eine Abbildung* $\mathfrak{f}$ *von* $\prod_{0 \leq i \leq p} I_i \times \prod_{0 \leq i \leq p} H_i$ *in* E verschärft Lipschitzsch (*bzw.* verschärft lokal Lipschitzsch) *ist, falls erstens* $\tilde{H}$ *nicht leer, zweitens* $\tilde{\mathfrak{f}}$ *eine Funktion und drittens* $\tilde{\mathfrak{f}}$ *Lipschitzsch (bzw. lokal Lipschitzsch) bezüglich der Norm* $[\tilde{\mathfrak{x}}]$ *ist. Wenn* $\mathfrak{f}$ *verschärft Lipschitzsch (bzw. verschärft lokal Lipschitzsch) in* $\prod_{0 \leq i \leq p} I_i \times \prod_{0 \leq i \leq p} H_i$ *ist, heißt die Gleichung* $D_0 \mathfrak{x} = \mathfrak{f}(t_0, \dots, t_p, \mathfrak{x}, D_1 \mathfrak{x}, \dots, D_p \mathfrak{x})$ *verschärft Lipschitzsch (bzw. verschärft lokal Lipschitzsch) in* $\prod_{0 \leq i \leq p} I_i \times \prod_{0 \leq i \leq p} H_i$.

Die zweite Bedingung besagt, daß für jeden Punkt t_0 von I_0 und für jede Funktion $\tilde{\mathfrak{x}}$ von $\tilde{H}$ die Funktion $\tau \to \mathfrak{f}(t_0, \tau, \tilde{\mathfrak{x}}(\tau), D_1 \tilde{\mathfrak{x}}(\tau), \dots, D_p \tilde{\mathfrak{x}}(\tau))$ zu $\tilde{E}$ gehört. Sie muß also insbesondere unendlich oft differenzierbar sein, was man in allgemeinen normierten Räumen der Funktion $\mathfrak{f}$ nicht ohne weiteres ansieht. Wie man in vielen Fällen auf konstruktivem Wege zu derartigen Funktionen gelangt, die überdies noch verschärft Lipschitzsch sind, lehrt der folgende Satz.

Satz 6. (a) *Mit den Funktionen* $\mathfrak{f}$ *und* $\mathfrak{g}$ *und für* $\alpha \in \mathbf{R}$ *sind auch die Funktionen* $\mathfrak{f} + \mathfrak{g}$ *und* $\alpha \cdot \mathfrak{f}$ *verschärft Lipschitzsch in* $\prod_{0 \leq i \leq p} I_i \times \prod_{0 \leq i \leq p} H_i$.

(b) *Ist* $\mathfrak{U}$ *eine stetige lineare Abbildung von* E *in* E, *dann ist mit* $\mathfrak{f}$ *auch die Funktion* $\mathfrak{U} \circ \mathfrak{f}$ *verschärft Lipschitzsch in* $\prod_{0 \leq i \leq p} I_i \times \prod_{0 \leq i \leq p} H_i$.

(a) Zunächst ist $\widetilde{H}$ nicht leer. Ferner beachte man, daß mit $\tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})$ und $\tilde{\mathfrak{g}}(t_0, \tilde{\mathfrak{x}})$ auch $\tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}}) + \tilde{\mathfrak{g}}(t_0, \tilde{\mathfrak{x}})$ und $\alpha \cdot \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})$ zu $\widetilde{E}$ gehören (Satz 1) und mit $t_0 \to \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})$ und $t_0 \to \tilde{\mathfrak{g}}(t_0, \tilde{\mathfrak{x}})$ auch $t_0 \to \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}}) + \tilde{\mathfrak{g}}(t_0, \tilde{\mathfrak{x}})$ und $t_0 \to \alpha \cdot \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})$ Regelfunktionen sind. Schließlich gelten nach Satz 1 die Ungleichungen

$$[(\tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}}) + \tilde{\mathfrak{g}}(t_0, \tilde{\mathfrak{g}})) - (\tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{y}}) + \tilde{\mathfrak{g}}(t_0, \tilde{\mathfrak{y}}))] \leq (k_1 + k_2)\,[\tilde{\mathfrak{x}} - \tilde{\mathfrak{y}}],$$

$$[\alpha \cdot \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}}) - \alpha \cdot \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{y}})] \leq |\alpha| \cdot k \cdot [\tilde{\mathfrak{x}} - \tilde{\mathfrak{y}}],$$

wo k_1, k_2, k LIPSCHITZ-Konstanten sind.

(b) Wiederum ist $\widetilde{H}$ nicht leer. Ferner bedenke man, daß mit $\tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})$ auch die Funktion $\mathsf{U} \circ \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})$ zu $\widetilde{E}$ gehört (Satz 2) und mit $t_0 \to \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})$ wegen der Stetigkeit von U auch $t_0 \to \mathsf{U} \circ \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})$ Regelfunktion ist. Endlich gilt nach Satz 2 die Ungleichung

$$[\mathsf{U} \circ \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}}) - \mathsf{U} \circ \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{y}})] \leq \|\mathsf{U}\| \cdot k \cdot [\tilde{\mathfrak{x}} - \tilde{\mathfrak{y}}],$$

wo k eine LIPSCHITZ-Konstante ist.

Bemerkung. Der Satz gilt auch, wenn man die Werte „verschärft Lipschitzsch" überall durch „verschärft lokal Lipschitzsch" ersetzt.

Beispiele. 1. Gegeben sei eine Abbildung $\mathfrak{a}$ von $\prod_{0 \leq i \leq p} I_i$ in E mit $\mathfrak{a}_{t_0} \in \widetilde{E}$; ferner sei $t_0 \to \mathfrak{a}_{t_0}$ Regelfunktion. Dann ist die Funktion

$$(t_0, \ldots, t_p, \mathfrak{x}_0, \ldots, \mathfrak{x}_p) \to \mathfrak{a}(t_0, \ldots, t_p)$$

verschärft Lipschitzsch in $\prod_{0 \leq i \leq p} I_i \times E^{p+1}$. In der Tat ist zunächst $\widetilde{H}$ nicht leer. Wegen $\tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}}) = \mathfrak{a}_{t_0}$ ist mit $\mathfrak{a}_{t_0}$ auch $\tilde{\mathfrak{f}}(t_0, \mathfrak{x})$ zu $\widetilde{E}$ gehörig und mit $t_0 \to \mathfrak{a}_{t_0}$ auch $t_0 \to \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})$ Regelfunktion. Die LIPSCHITZ-Bedingung ist wegen $[\tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}}) - \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{y}})] = 0$ in trivialer Weise erfüllt.

2. Für alle $0 \leq i \leq p$ ist die Funktion $(t_0, \ldots, t_p, \mathfrak{x}_0, \ldots, \mathfrak{x}_p) \to \mathfrak{x}_i$ verschärft Lipschitzsch in $\mathsf{R}^{p+1} \times E^{p+1}$. Wiederum ist $\widetilde{H}$ nicht leer. Wegen $\tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}}) = \tilde{\mathfrak{x}}$ (bzw. $D_j \tilde{\mathfrak{x}}, j \geq 1$) gehört (nach Satz 2) mit $\tilde{\mathfrak{x}}$ auch $\tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})$ zu $\widetilde{E}$; ferner ist $t_0 \to \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})$ konstant, also Regelfunktion. Die LIPSCHITZ-Bedingung ist wegen $[\tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}}) - \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{y}})] \leq [\tilde{\mathfrak{x}} - \tilde{\mathfrak{y}}]$ mit der Konstanten 1 erfüllt.

3. Für jede Familie $(a_i)_{0 \leq i \leq p}$ von reellen Zahlen ist nach obigem und Satz 5 die Funktion $(t_0, \ldots, t_p, \mathfrak{x}_0, \ldots, \mathfrak{x}_p) \to \sum_{0 \leq i \leq p} a_i \mathfrak{x}_i$ verschärft Lipschitzsch in $\mathsf{R}^{p+1} \times E^{p+1}$ (vgl. 1., Beispiel 1).

4. Sei $p = 1$, $E = \mathsf{C}^3$, a eine Abbildung von $I_0 \times I_1$ in C mit $a_{t_0} \in \widetilde{\mathsf{C}}$; ferner sei $t_0 \to a_{t_0}$ Regelfunktion. Dann ist die Funktion

$$(t_0, t_1, (\xi_{0\varrho})_{1 \leq \varrho \leq 3}, \quad (\xi_{1\varrho})_{1 \leq \varrho \leq 3}) \to (\xi_{02}, \xi_{13} + a(t_0, t_1), \xi_{12})$$

verschärft Lipschitzsch in $(I_0 \times I_1) \times (\mathsf{C}^3 \times \mathsf{C}^3)$; denn es sind die Abbildungen

$$(t_0, t_1, (\xi_{0\varrho}), (\xi_{1\varrho})) \to (0, a(t_0, t_1), 0)$$

$$(t_0, t_1, (\xi_{0\varrho}), (\xi_{1\varrho})) \to (\xi_{01}, \xi_{02}, \xi_{03}) \to (\xi_{02}, 0, 0)$$

$$(t_0, t_1, (\xi_{0\varrho}), (\xi_{1\varrho})) \to (\xi_{11}, \xi_{12}, \xi_{13}) \to (0, \xi_{13}, \xi_{12})$$

nach obigem und Satz 5 verschärft Lipschitzsch, also auch deren Summe (vgl. § 1, Beispiel 2).

5. Die Funktion $(t_0, t_1, x_0, x_1) \to t_1 \cdot x_0$ ist nicht lokal verschärft Lipschitzsch; denn es ist $D_1^{n_1} \tilde{f}(t_0, \tilde{x}) = \varphi \cdot D_1^{n_1} \tilde{x} + n \cdot D_1^{n-1} \tilde{x}$, wobei φ die kanonische Abbildung von I_1 in $\mathbf{C}$ bedeutet. Wählen wir für $\tilde{x}$ speziell die (zu $\tilde{E}$) gehörige Funktion $t_1 \to e^{t_1}$, so geht $\tilde{f}(t_0, \tilde{x})$ in die Funktion

$$t_1 \to t_1 e^{t_1} + n e^{t_1} \geq (1+n) \cdot e$$

über. Wegen $\sup_{n \in \mathbf{N}} (1+n) \cdot e = \infty$ ist also $\tilde{f}$ keine Funktion (vgl. § 3, Bemerkung).

§ 5. Existenzsätze

Der folgende Satz ist eine wörtliche Übertragung des entsprechenden, für gewöhnliche Differentialgleichungen geltenden Satzes (vgl. [4]).

Hauptsatz. *Seien $(I_i)_{0 \leq i \leq p}$ eine Familie von nicht auf einen Punkt reduzierten, in $\mathbf{R}$ enthaltenen Intervallen, $t_* \in I_0$ der Anfang von I_0, $(H_i)_{0 \leq i \leq p}$ eine Familie von offenen, nicht leeren Teilen von E, $\mathfrak{f}$ eine in $\prod_{0 \leq i \leq p} I_i \times \prod_{0 \leq i \leq p} H_i$ verschärft lokal* LIPSCHITZsche *Funktion.*

(a) *Zu jedem $\tilde{x}_* \in \tilde{H}$ existiert ein größtes Intervall $J \subset I_0$ mit dem Anfang $t_* \in J$ derart, daß in dem Produkt $\prod_{0 \leq i \leq p} I_i'$ mit $I_0' = J$ und $I_j' = I_j$ $(1 \leq j \leq p)$ eine starke Lösung $\mathfrak{u}$ der Gleichung $D_0 \mathfrak{x} = \mathfrak{f}(t_0, \ldots, t_p, \mathfrak{x}, D_1 \mathfrak{x}, \ldots, D_p \mathfrak{x})$ existiert, die (bzw. deren Ableitung $D_j \mathfrak{u}$, $1 \leq j \leq p$) ihre Werte in H_0 (bzw. H_j) annimmt und in die Funktion $\tilde{x}_*$ im Punkt t_* übergeht; diese Lösung ist eindeutig.*

(b) *Wenn $J \neq J_0$, dann ist J ein halb-offenes Intervall $\langle t^*, \beta \langle$ von endlicher Länge; überdies ist für jeden kompakten Teil $K \subset \tilde{H}$ auch das reziproke Bild durch $t_0 \to \mathfrak{u}_{t_0}$ von K ein kompakter Teil von $\mathbf{R}$.*

(c) *Wenn J beschränkt ist, und wenn $\tilde{\mathfrak{f}}(t_0, \mathfrak{u}_{t_0})$ in J beschränkt ist, dann hat $t_0 \to \mathfrak{u}_{t_0}$ einen linksseitigen Grenzwert $\check{c}$ im Endpunkt von J; wenn überdies $J \neq I_0$, dann ist $\check{c}$ ein Begrenzungspunkt von $\tilde{H}$ in $\tilde{E}$.*

Auf $\tilde{\mathfrak{f}}$ umgeschrieben, lauten die Voraussetzungen: Seien $I_0 \subset \mathbf{R}$ ein (nicht auf einen Punkt reduziertes) Intervall, $t_* \in I_0$ der Anfang von I_0, $\tilde{H}$ ein offener, nicht leerer Teil von $\tilde{E}$ (Satz 3 und Definition 2), $\tilde{\mathfrak{f}}$ eine in $I_0 \times \tilde{H}$ lokal LIPSCHITZsche Funktion (Definition 2). Daraus folgt die Existenz eines Intervalls $J \subsetneq I_0$ und einer Lösung $\tilde{\mathfrak{u}}$ der Gleichung $\tilde{\mathfrak{x}}' = \tilde{\mathfrak{f}}(t_0, \tilde{\mathfrak{x}})$, die die Eigenschaften (a), (b), (c) – sinngemäß abgewandelt – des vorstehenden Satzes besitzt. Wir bilden eine in $\prod_{0 \leq i \leq p} I_i'$ erklärte Funktion $\mathfrak{u}$ mit Werten in E, indem wir jedem Punkt (t_0, τ) des Produktes $\prod_{0 \leq i \leq p} I_i'$ den Wert der Funktion $\tilde{\mathfrak{u}}(t_0)$ für den Punkt τ zuordnen. Offensichtlich stimmt $t_0 \to \mathfrak{u}_{t_0}$ mit $\tilde{\mathfrak{u}}$ überein, woraus sich zusammen mit Satz 4 und Definition 1 die Behauptung ergibt.

Bemerkung. Man wird zunächst die Verwendung der mit einer Tilde versehenen Größen $\tilde{E}$, $\tilde{H}$, $\check{c}$ in der Formulierung des Satzes als problemfremd empfinden. Überlegt man sich jedoch, daß die Lösungsfläche durch gleich-

mäßiges Fortschreiten längs der I_0-Achse auf der ganzen Breite der Q-Achse entsteht, so erscheint diese Auffassung als die naturgegebene. (Siehe Fig. 1!)

Beispiele. 1. Wir betrachten die lineare homogene Gleichung $D_0 x = a x + b \cdot D_1 x$ mit konstanten Koeffizienten. Sie ist verschärft lokal Lipschitzsch (vgl. § 4, Beispiel 3). Die zugehörige gewöhnliche Gleichung lautet $\tilde{x}' = a\tilde{x} + b \cdot D_1 \tilde{x}$ oder $\tilde{x}' = A \cdot \tilde{x}$, wobei A die stetige lineare Abbildung $\tilde{x} \to a\tilde{x} + b \cdot D_1 \tilde{x}$ bedeutet (vgl. § 1, Beispiel 1). Wir setzen die Lösung $\tilde{u}$ in der Form $t_0 \to C(t_0, t_*) \cdot \tilde{x}_*$ an; $C(t_0, t_*)$ nennt man die Resolvente der linearen Gleichung. Nach bekannten Sätzen ist $C(t_0, t_*) = \exp\big(A \cdot (t_0 - t_*)\big)$, also $\tilde{u}(t_0) = \exp\big(A \cdot (t_0 - t_*)\big) \cdot \tilde{x}_* = \sum_{n \in \mathbf{N}} (A^n \cdot \tilde{x}_*) \frac{(t_0 - t_*)^n}{n!}$. Wir berechnen:

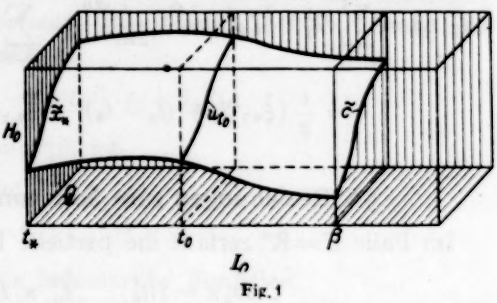


Fig. 1

$$A^n \tilde{x}_* = \sum_{0 \le \nu \le n} \binom{n}{\nu} a^\nu \cdot D_1^{n-\nu} \tilde{x}_* \cdot b^{n-\nu},$$

also

$$\tilde{u}(t_0) = \sum_{\nu \in \mathbf{N}} \frac{1}{\nu!} a^\nu (t_0 - t_*)^\nu \cdot \sum_{\mu \in \mathbf{N}} \frac{1}{\mu!} b^\mu (t_0 - t_*)^\mu \cdot D_1^\mu \tilde{x}_*$$

und schließlich, wenn wir zu u übergehen,

$$u(t_0, t_1) = e^{a(t_0 - t_*)} \cdot \tilde{x}_*\big(t_1 + b(t_0 - t_*)\big).$$

2. Wir befassen uns jetzt mit der hyperbolischen Gleichung

$$\frac{\partial^2 z}{\partial x^2} - \frac{\partial^2 z}{\partial y^2} = a(x, y),$$

in unserer Schreibweise

$$D_0 \xi_1 = \xi_2, \quad D_0 \xi_2 = D_1 \xi_3 + a(t_0, t_1), \quad D_0 \xi_3 = D_1 \xi_2.$$

Sie ist verschärft lokal Lipschitzsch (vgl. § 4, Beispiel 4). Die zugehörige gewöhnliche Gleichung lautet

$$\tilde{\xi}_1' = \tilde{\xi}_2, \quad \tilde{\xi}_2' = D_1 \tilde{\xi}_3 + a_{t_0}, \quad \tilde{\xi}_3' = D_1 \tilde{\xi}_2$$

oder $(\tilde{\xi}_\varrho')_{1 \le \varrho \le 3} = A \cdot (\tilde{\xi}_\varrho)_{1 \le \varrho \le 3} + \mathfrak{b}(t_0)$, wobei A die stetige homogene lineare Abbildung $(\tilde{\xi}_\varrho)_{1 \le \varrho \le 3} \to (\tilde{\xi}_2, D_1 \tilde{\xi}_3, D_1 \tilde{\xi}_2)$, $\mathfrak{b}(t_0)$ die Funktion $t_1 \to \big(0, a(t_0, t_1), 0\big)$ bedeuten (vgl. § 1, Beispiel 2). Wir setzen die Lösung $\tilde{\mathfrak{u}}$ der homogenen Gleichung in der Form $t_0 \to C(t_0, t_*) \cdot \tilde{\mathfrak{x}}_*$ an, wo wiederum $C(t_0, t_*)$ die Resolvente der linearen Gleichung ist. Wie bekannt, läßt sich $C(t_0, t_*)$ in der Gestalt $\exp\big(A \cdot (t_0 - t_*)\big)$, also $\tilde{\mathfrak{u}}(t_0)$ in der Gestalt $\exp\big(A \cdot (t_0 - t_*)\big) \cdot \tilde{\mathfrak{x}}_* = \sum_{n \in \mathbf{N}} (A^n \cdot \tilde{\mathfrak{x}}_*) \frac{(t_0 - t_*)^n}{n!}$ schreiben. Die Berechnung ergibt

$$A^{2n} \cdot \tilde{\mathfrak{x}}_* = (D_1^{2n-1} \tilde{\xi}_{*3}, D_1^{2n} \tilde{\xi}_{*2}, D_1^{2n} \tilde{\xi}_{*3}) \quad (n \in \mathbf{N}^*),$$

$$A^{2n+1} \cdot \tilde{\mathfrak{x}}_* = (D_1^{2n} \tilde{\xi}_{*2}, D_1^{2n+1} \tilde{\xi}_{*3}, D_1^{2n+1} \tilde{\xi}_{*2}) \quad (n \in \mathbf{N}).$$

Beschränken wir uns auf die allein interessierende erste Komponente und bezeichnen ihren Wert für den Punkt (t_0, t_1) (ohne Verwechslungen befürchten zu müssen) mit $u(t_0, t_1)$, so kommt unter Beachtung von $\tilde{\xi}_{*2} = D_1 \tilde{\xi}_{*1}$ schließlich

$$u(t_0, t_1) = \sum_{n \in \mathbf{N}} D_1^{2n} \cdot \tilde{\xi}_{*1}(t_1) \cdot \frac{(t_0 - t_*)^{2n}}{(2n)!} + \sum_{n \in \mathbf{N}} D_1^{2n} \tilde{\xi}_{*2}(t_1) \frac{(t_0 - t_*)^{2n+1}}{(2n+1)!}$$

$$= \frac{1}{2} \left(\tilde{\xi}_{*1}(t_1 + (t_0 - t_*)) + \tilde{\xi}_{*1}(t_1 - (t_0 - t_*)) \right) + \frac{1}{2} \int_{t_1 - (t_0 - t_*)}^{t_1 + (t_0 - t_*)} \tilde{\xi}_{*2}(\tau)\, d\tau .$$

§ 6. Beziehungen zum Satz von CAUCHY-KOWALEWSKI

Im Falle $E = \mathbf{R}^n$ zerfällt die partielle Differentialgleichung

$$D_0 \mathbf{x} = \mathbf{f}(t_0, \ldots, t_p, \mathbf{x}, D_1 \mathbf{x}, \ldots, D_p \mathbf{x})$$

in ein endliches Normalsystem

$$D_0 \xi_\varrho = f_\varrho(t_0, \ldots, t_p, (\xi_\sigma)_{1 \le \sigma \le n}, (D_1 \xi_\sigma)_{1 \le \sigma \le n}, \ldots, (D_p \xi_\sigma)_{1 \le \sigma \le n}) \quad (1 \le \varrho \le n)$$

von skalaren partiellen Differentialgleichungen, das man unter geeigneten Voraussetzungen nach dem CAUCHY-KOWALEWSKIschen Verfahren (vgl. [2] und [5]) lösen kann. Diese bestehen wesentlich in der Holomorphie der rechten Seiten des obigen Gleichungssystems und der Anfangswerte, während für die Anwendung unseres Verfahrens die rechten Seiten einer verschärften LIPSCHITZ-Bedingung genügen und die Anfangswerte unendlich oft differenzierbar und ableitungsbeschränkt sein müssen. Dies bedeutet zwar ebenfalls die Holomorphie der Anfangswerte und (in gewissem Sinne) der rechten Seiten des Gleichungssystems, noch dazu mit einer sehr einschneidenden Wachstumsbeschränkung der Koeffizienten der Potenzreihen, aber nur hinsichtlich der Variablen $t_1, \ldots, t_p$; bezüglich t_0 genügt es, Regelfunktionen vorauszusetzen.

Die unterschiedlichen Hypothesen bewirken naturgemäß unterschiedliche Eigenschaften der Lösungen. Der Satz von CAUCHY-KOWALEWSKI garantiert von vornherein die Existenz einer holomorphen Lösung bezüglich aller Variablen $t_0, \ldots, t_p$, während der Hauptsatz dieser Arbeit die Holomorphie, genauer sogar die unbeschränkte Differenzierbarkeit und Ableitungsbeschränktheit bezüglich der Variablen $t_1, \ldots, t_p$, hinsichtlich der Variablen t_0 jedoch nur die Differenzierbarkeit auf dem Komplement einer abzählbaren Menge (Stammfunktion einer Regelfunktion) sicherstellt. Trotz der etwa bezüglich t_0 auftretenden Knickstellen bleibt dabei in bemerkenswerter Weise die Eindeutigkeit — solange es sich um starke Lösungen handelt — gewährleistet. Es verdient noch hervorgehoben zu werden, daß nach CAUCHY-KOWALEWSKI die Existenz der Lösung nur in einer Umgebung eines Punktes $(t_0, \ldots, t_p)$ nachgewiesen wird, während die Lösung nach unserem Verfahren auf der gesamten Breite Q eines p-dimensionalen Intervalls in einer Umgebung des Anfangspunktes t_* erklärt werden kann.

Beispiele. 1. Die rechte Seite der Gleichung

$$D_0 x = |t_0| \cdot t_1$$

ist eine in $\mathbf{R}^4$ erklärte, verschärft LIPSCHITZsche, jedoch nicht holomorphe Funktion $(t_0, t_1, x_0, x_1) \to |t_0| \cdot t_1$ mit Werten in $\mathbf{R}$. Während der Existenzbeweis von CAUCHY-KOWALEWSKI versagt, liefert unser Satz als Lösung die im Punkt t_* in $\tilde{x}_*$ übergehende Funktion

$$(t_0, t_1) \to \tfrac{1}{2}(t_0|t_0| - t_*|t_*|) \cdot t_1 + \tilde{x}_*(t_1),$$

die allerdings nur bezüglich t_1 holomorph ist.

2. Die rechte Seite der Gleichung

$$D_0 x = t_0(1 - t_1)^{\frac{1}{2}}$$

ist eine in $\mathbf{R} \times]-1, 1[\times \mathbf{R}^2$ erklärte holomorphe Funktion

$$(t_0, t_1, x_0, x_1) \to t_0(1 - t_1)^{\frac{1}{2}}$$

mit Werten in $\mathbf{R}$, die wegen

$$\left|(D^n)_0^{t_1}(1 - t_1)^{\frac{1}{2}}\right| \geq (n-1)!/2^n \qquad (n \geq 1)$$

in keinem Punkte von der Form $(t_0, 0, x_0, x_1)$ enthaltenden Teil von $\mathbf{R} \times]-1, 1[\times \mathbf{R}^2$ verschärft lokal Lipschitzsch ist. Während unser Existenzbeweis versagt, liefert der CAUCHY-KOWALEWSKIsche Satz als Lösung die im Punkt t_* in $\tilde{x}_*$ übergehende holomorphe Funktion

$$(t_0, t_1) \to \tfrac{1}{2}(t_0^2 - t_*^2) \cdot (1 - t_1)^{\frac{1}{2}} + \tilde{x}_*(t_1).$$

3. Die rechte Seite der Gleichung

$$D_0 x = t_0 t_1$$

ist eine in $\mathbf{R}^4$ erklärte holomorphe, verschärft LIPSCHITZsche Funktion $(t_0, t_1, x_0, x_1) \to t_0 t_1$ mit Werten in $\mathbf{R}$. Sowohl der CAUCHY-KOWALEWSKIsche als auch unser Satz liefern als Lösung die im Punkt t_* in $\tilde{x}_*$ übergehende holomorphe Funktion

$$(t_0, t_1) \to \tfrac{1}{2}(t_0^2 - t_*^2) \cdot t_1 + \tilde{x}_*(t_1),$$

falls $\tilde{x}_*$ holomorph (bei CAUCHY-KOWALEWSKI) bzw. unendlich oft differenzierbar und ableitungsbeschränkt (bei uns) ist.

Literatur

[1] WAŻEWSKI, T.: Über die Bedingungen der Existenz der Integrale partieller Differentialgleichungen erster Ordnung. Math. Z. **43** (1938). — [2] GOURSAT, E.: Cours d'analyse mathématique, tome II, p. 360. Paris 1942. — [3] BOURBAKI, N.: Topologie générale, chap. IX, X. Paris 1948 bzw. 1949. — [4] BOURBAKI, N.: Fonctions d'une variable réelle, chap. IV. Paris 1951. — [5] SAUER, R.: Anfangswertprobleme bei partiellen Differentialgleichungen. Berlin-Göttingen-Heidelberg 1952.

München, Mathematisches Institut der Universität

(Eingegangen am 10. Februar 1956)

PETERSEN, G. M.
Math. Zeitschr. Bd. 65, S. 494f–496 (1956)

Inclusion between limitation methods

By

GORDON M. PETERSEN

We shall write $B > A$ when B is a limitation method that is strictly stronger than A, that is B limits all bounded sequences that are A limitable and also limits at least one bounded sequence that is not A limitable. Our purpose in this note is to show that if a regular matrix $B = (b_{m\,n})$ is strictly stronger than a regular matrix method $A = (a_{m\,n})$, then there exists a regular matrix $C = (c_{m\,n})$ such that B is strictly stronger than C and C is strictly stronger than A, i.e. $B > C > A$.

By a submethod of a regular matrix, $A = (a_{m\,n})$, we mean the regular matrix formed by using any infinite set of rows taken from the original matrix. It is evident, for example, that any bounded sequence that is not A limitable, is limited by at least one of its submethods.

The result we have stated was originally obtained with a different proof by BRUDNO [1]. Similar methods of constructing certain B limitable bounded sequences were developed by AGNEW [2] and ZELLER [3]. Submethods of a regular matrix limitation method have been studied in [4].

This result can not be generalized to state that if B is stronger than A and limits at least one sequence that is not A limitable, then there is a matrix C such that B is stronger than C and limits a non C limitable sequence and C is stronger than A and limits a non A limitable sequence. The author has shown [5], that the limitation methods described by

$$t_n = 2s_{n-1} - s_n$$

and

$$\tau_n = 4s_{n-2} - 4s_{n-1} + s_n$$

limit only sequences of the form $C_1 \cdot 2^n + s'_n$ and $C_2 \cdot n \cdot 2^n + C_3 \cdot 2^n + s''_n$ respectively where C_1, C_2, and C_3 are arbitrary constants and $\{s'_n\}$ and $\{s''_n\}$ are arbitrary convergent sequences. Clearly there is no matrix $C = (c_{m\,n})$ that satisfies the conditions listed above.

We first prove a lemma.

LEMMA. *If B is strictly stronger than A, then there is a submethod of A that limits some, but not all, of the bounded sequences that are B limitable.*

PROOF. Suppose that the A and B transforms are given by

$$t_m = \sum_{n=1}^{\infty} a_{m\,n} s_n \quad \text{and} \quad \tau_m = \sum_{n=1}^{\infty} b_{m\,n} s_n .$$

We can assume that the bounded sequence $\{s_n^1\}$ which is B limitable but not A limitable is limited by B to 0 and that $\limsup t_m^1 = u$ and $\liminf t_m^1 = 1$. Let $\sup_n |s_n| = H$ and $\sup_m \sum_{n=1}^{\infty} |b_{m\,n}|\,|s_n| = M$.

We shall consider the partial sums $\{u_k\}$ of the series $1-1+0+\cdots+0+\frac{1}{2}+\frac{1}{2}+0+\cdots+0-\frac{1}{2}-\frac{1}{2}+0+\cdots+0+\cdots+1/n+\cdots+1/n+0+\cdots+0-1/n-\cdots-1/n+0+\cdots+0+\cdots$ where $1/n$ appears in the series $2n$ times, the first n times with the plus sign followed by $g(n)$ 0's; the second n times with the minus sign followed by $h(n)$ 0's, $g(n)$ and $h(n)$ being integral valued functions of n. It is evident that $0 \leq u_k \leq 1$ for every k and if $\eta_k = |u_{k+1} - u_k|$, then $\lim_{k\to\infty} \eta_k = 0$.

Since A and B are regular, there is a sequence $\{m_k\}$ and two non decreasing functions $\lambda(m)$ and $\varkappa(m)$ such that

$$\sum_{n=1}^{\lambda(m)} |a_{m\,n}| < \varepsilon_m, \quad \sum_{n=1}^{\lambda(m)} |b_{m\,n}| < \varepsilon_m, \quad \sum_{n=\varkappa(m)}^{\infty} |a_{m\,n}| < \varepsilon_m, \quad \sum_{n=\varkappa(m)}^{\infty} |b_{m\,n}| < \varepsilon_m$$

where $\lambda(m_k) = \varkappa(m_{k-1})$, $k = 2, 3, \ldots$ and $\lim_{m\to\infty} \varepsilon_m = 0$. If necessary, we may take $\lambda(m)$ so that $\lambda(m+1) - \lambda(m) \leq 1$ for every m. Since $B - \lim s_n^1 = 0$, it is evident that if $\gamma_m = \left| \sum_{\lambda(m)}^{\varkappa(m)} b_{m\,n} s_n^1 \right|$, then $\lim_{m\to\infty} \gamma_m = 0$.

We now define the sequence $\{s_n^2\}$. Let $s_n^2 = s_n^1$ for $n \leq \lambda(m_1)$ and let $s_n^2 = u_k\, s_n^1$ whenever $\lambda(m_k) < n \leq \varkappa(m_k)$. It is clear that $\{s_n^2\}$ depends on the choice of $g(n)$ and $h(n)$; in any case we shall now show that $B - \lim s_n^2 = 0$.

Suppose that $m_k < \mu \leq m_{k+1}$,

$$\tau_\mu^2 = \sum_{n=1}^{\infty} b_{\mu\,n}\, s_n^2 = \sum_{n=1}^{\lambda(\mu)-1} b_{\mu\,n}\, s_n^2 + u_k \sum_{n=\lambda(\mu)}^{\varkappa(\mu)} b_{\mu\,n}\, s_n^1 + (u_{k+1} - u_k) \sum_{n=\varkappa(m_k)}^{\varkappa(\mu)} b_{\mu\,n}\, s_n^1 + \sum_{n=\varkappa(\mu)+1}^{\infty} b_{\mu\,n}\, s_n^2,$$

since $\varkappa(\mu) \leq \varkappa(m_{k+1})$ and $\lambda(m_k) \leq \lambda(\mu) \leq \varkappa(m_k)$. It follows that

$$|\tau_\mu^2| \leq \sum_{n=1}^{\lambda(\mu)-1} |b_{\mu\,n}|\,|s_n^2| + |u_k| \left| \sum_{n=\lambda(\mu)}^{\varkappa(\mu)} b_{\mu\,n}\, s_n^1 \right| + |u_{k+1} - u_k| \sum_{n=\varkappa(m_k)}^{\varkappa(\mu)} |b_{\mu\,n}|\,|s_n^1| + \\ + \sum_{n=\varkappa(\mu)+1}^{\infty} |b_{\mu\,n}|\,|s_n^2|,$$

$$\leq H\,\varepsilon_\mu + \gamma_\mu + M\,\eta_k + H\,\varepsilon_\mu,$$

so that $B - \lim s_n^2 = 0$.

On the other hand, let the sequences $\{m_i\}$ and $\{m_j\}$, $t_{m_i} = t_i$, $t_{m_j} = t_j$, have the property $\lim_{i\to\infty} t_i^1 = u$ and $\lim_{j\to\infty} t_j^1 = 1$. Clearly we can choose $g(n)$ and $h(n)$ so that

$$s_n^2 = s_n^1 \quad \text{for all} \quad \lambda(m_i) < n \leq \varkappa(m_i),$$

$$s_n^2 = s_n^1 \quad \text{for all} \quad \lambda(m_j) < n \leq \varkappa(m_j),$$

over infinite sets of i and j, $\{\bar{\bar{i}}\}$ and $\{\bar{\bar{j}}\}$:

$$s_n^2 = 0 \quad \text{for all} \quad \lambda(m_i) < n \leq \varkappa(m_i),$$
$$s_n^2 = 0 \quad \text{for all} \quad \lambda(m_j) < n \leq \varkappa(m_j)$$

over infinite sets of i and j, $\{\bar{i}\}$ and $\{\bar{j}\}$.

This choice means that either $\lim t_{\bar{i}}^2 \neq \lim t_{\bar{i}}^1$ or $\lim t_{\bar{\bar{j}}}^2 \neq \lim t_{\bar{j}}^2$ so that $\{s_n^2\}$ is not A limitable. However a submethod A' consisting of infinitely many $t_{\bar{i}}$ and $t_{\bar{j}}$ will limit $\{s_n^2\}$ to 0; it will not limit $\{s_n^1\}$ since $\lim t_{\bar{i}}^1 \neq \lim t_{\bar{j}}^1$.

We are now able to prove our theorem.

THEOREM. *If $B > A$, there is a C, such that $B > C > A$.*

PROOF. We define our new matrix $C = (c_{m\,n})$ by intertwining alternately the rows of A' as described in the lemma and B. The matrix C is the required matrix, it limits all A limitable sequences (because A and B are compatible, see BRUDNO [1]) and does not limit any sequences that are not B limitable; it limits $\{s_n^2\}$ to 0 and so is stronger than A, but it does not limit $\{s_n^1\}$ and so is weaker than B.

References

[1] BRUDNO, A.: Summation of Bounded Sequences by Matrices. Rec. Math. (Mat. Sbornik) N.S. **16**, (58) 191—243 (1945) (in Russian, English summary). — [2] AGNEW, R. P.: Convergence fields of methods of summability. Annals of Math. **46**, 93—101 (1945). — [3] ZELLER, K.: Faktorfolgen bei Limitierungsverfahren. Math. Z. **56**, 134—151 (1952). — [4] GOFFMAN, CASPER and G. M. PETERSEN: Submethods of Regular Matrix Summability Methods. Canad. J. Math. **8**, 40—46 (1956). — [5] PETERSEN, G. M.: A Note on Divergent Series. Canad. J. Math. **4**, 445—454 (1952). — [6] MAZUR, S., and W. ORLICZ: Sur les methodes linéaires de sommation. C. R. Acad. Sci. Paris **196**, 32—34 (1933). — [7] MAZUR, S., and W. ORLICZ: On linear methods of summability. Studia Math. **14**, 129—160 (1955).

University College, Swansea, Wales (England)

(Eingegangen am 22. Februar 1956)